UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE GENÉTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA

LUDMILA ARRUDA DE ASSIS

Caracterização das interações moleculares e mRNAs alvos de proteínas com domínio de ligação ao RNA implicadas no processo de tradução em *Leishmania infantum*

Recife
2019

LUDMILA ARRUDA DE ASSIS

Caracterização das interações moleculares e mRNAs alvos de proteínas com domínio de ligação ao RNA implicadas no processo de tradução em *Leishmania infantum*

Tese apresentada ao Programa de Pós-Graduação em Genética da Universidade Federal de Pernambuco como parte dos requisitos exigidos para obtenção do título de Doutor em Genética.

Área de concentração: Biologia Molecular

Orientador: Osvaldo Pompílio de Melo Neto
Coorientador (a): Tamara De Carli da Costa Lima

Recife
2019

Catalogação na fonte
Elaine C Barroso (CRB4/1728)

Assis, Ludmila Arruda de

Caracterização das interações moleculares e mRNAs alvos de proteínas com domínio de ligação ao RNA implicadas no processo de tradução em *Leishmania infantum* / Ludmila Arruda de Assis- 2019.

159 folhas: il., fig., tab.
Orientador: Osvaldo Pompílio de Melo Neto
Coorientadora: Tamara de Carli da Costa Lima

Tese (doutorado) – Universidade Federal de Pernambuco. Centro de Biociências. Programa de Pós-Graduação em Genética. Recife, 2019.
Inclui referências e apêndice

1. Expressão gênica 2. Tripanossomatídeos 3. Leishmania I. Melo Neto, Osvaldo Pompílio de (orient.) II. Lima, Tamara de Carli da Costa (coorient.) III. Título

572.865 CDD (22.ed.) UFPE/CB-2019-355

LUDMILA ARRUDA DE ASSIS

Caracterização das interações moleculares e mRNAs alvos de proteínas com domínio de ligação ao RNA implicadas no processo de tradução em *Leishmania infantum*

Tese apresentada ao Programa de Pós-Graduação em Genética da Universidade Federal de Pernambuco como parte dos requisitos exigidos para obtenção do título de Doutor em Genética.

Aprovado em: 23/07/2019

BANCA EXAMINADORA

_____

Osvaldo Pompílio de Melo Neto (Orientador)
Instituto Aggeu Magalhães/Fundação Oswaldo Cruz

_____

Marcos Antonio de Morais Junior (Examinador interno)
Universidade Federal de Pernambuco

_____

Danilo Elias Xavier (Examinador externo)
Instituto Aggeu Magalhães/Fundação Oswaldo Cruz

_____

Gabriel da Luz Wallau (Examinador externo)
Instituto Aggeu Magalhães/Fundação Oswaldo Cruz

_____

Flávia Figueira Aburjaile (Examinador externo)
Universidade Federal de Pernambuco

À minha mãe por estar sempre ao meu lado

AGRADECIMENTOS

Aos meus orientadores: Dr. Osvaldo Pompílio de Melo Neto, por todo incentivo e conhecimento dado e pela confiança em mim depositada para assumir esse projeto; Dr. Mark Carrington, por me receber em seu laboratório, pelos ensinamentos e pela paciência; e Dra. Tamara De Carli da Costa lima, pelas contribuições e amizade.

Aos órgãos de fomento CAPES e CNPq pelo apoio financeiro, ao Instituto Aggeu Magalhães (IAM-FIOCRUZ) pela infraestrutura e ao Programa de Pós-Graduação em Genética (PPGG-UFPE) pelo suporte acadêmico. Meu obrigada especial para os secretários Manassés da PPGG e Carol da Microbiologia pelo apoio administrativo.

A todo o grupo de pesquisa e aos amigos da Microbiologia por estarem sempre dispostos a ajudar, pelas boas conversas e risadas e pelo companheirismo diário, em especial para Adriana N., Allana, Artur, Camila, Carol, Christian, Dani, Eden, Gustavo, Irassandra, Isis, Maria, Moezio, Thaíse, Tallyta, Wagner e Yallen.

Aos colegas de Cambridge: Alex Klein, Camilla Trevor, Erick Eidy, Isobel Hambleton e Olivia Macleod, pelo auxilio no laboratório e pelos bons momentos fora dele.

As amigas e fadas sensatas Ray, Patrícia e Sheyla, ao meu namorado Philippe e toda minha família pelo apoio emocional, amor, carinho, dedicação e compreensão.

We keep moving forward, opening new doors, and doing new things, because we're curious and curiosity keeps leading us down new paths (WALT DISNEY, 2007).

RESUMO

A regulação da expressão gênica em *Leishmania* tem uma grande dependência do controle da estabilidade e tradução dos seus mRNAs. Em eucariotos, ambos são eventos associados a proteína de ligação a cauda poli-A (PABP) e outras proteínas com domínios de ligação a RNA, as RBPs. Em *Leishmania* foram descritas três PABPs, sendo que a PABP1 se associa a mRNAs alvos, complexos proteicos e RBPs distintas das PABP2/3. Seu papel funcional, contudo, ainda não está claro. Esse trabalho buscou auxiliar na definição de suas funções a partir da caracterização de RBPs parceiras, as RBP23, DRBD2 e ZC3H41. Extratos citoplasmáticos de *Leishmania infantum* expressando cada RBP fusionada com o peptídeo HA foram utilizados em ensaios de imunoprecipitação, com seus parceiros e mRNAs alvos identificados por espectrometria de massas e sequenciamento de RNAs. Foi demonstrada a interação preferencial da RBP23 com a PABP1 e o complexo de tradução EIF4G3/EIF4E4, da DRBD2 com as PABP2/PABP3 e da ZC3H41 com as três PABPs. Ensaios de interação proteína-proteína confirmaram uma ligação direta das RBP23 e DRBD2 com as três PABPs, mapeando regiões especificas das PABPs envolvidas nessa ligação. Os mRNAs co-precipitados com a RBP23, e em menor escala a ZC3H41, codificam preferencialmente proteínas ribossomais, semelhante a PABP1 e diferentemente da DRBD2. Estes dados sugerem um novo modelo para o reconhecimento destes mRNAs pelas RBP23 e PABP1 em *Leishmania*.

Palavras-chave: tripanossomatídeos. Expressão gênica. PABP. RBP

ABSTRACT

The regulation of gene expression in *Leishmania* depends largely on the stability and translation control of its mRNAs. In eukaryotes, both events are associated with poly-A tail binding protein (PABP) and other proteins with RNA binding domains, the RBPs. Three PABPs were described in *Leishmania*, and PABP1 is associated with target mRNAs, protein complexes and RBPs distinct from those associated to PABP2/3. However, its functional role is still unclear. This work aimed to define its functions from the characterization of partner RBPs, the RBP23, DRBD2 and ZC3H41. Cytoplasmic extracts of *Leishmania infantum* expressing each RBP fused to HA peptide were used in immunoprecipitation assays, with their partners and target mRNAs identified by mass spectrometry and RNA sequencing. It was demonstrated the preferential interaction of RBP23 with PABP1 and translation complex EIF4G3 / EIF4E4, DRBD2 with PABP2 / PABP3 and ZC3H41 with the three PABPs. Protein-protein interaction assays confirmed a direct binding of RBP23 and DRBD2 with the three PABPs, mapping specific regions of the PABPs involved in that binding. mRNAs co-precipitated with RBP23, and on a smaller scale with ZC3H41, preferentially encode ribosomal proteins, similar to PABP1 and unlike DRBD2. These data suggest a new model for the recognition of these mRNAs by RBP23 and PABP1 in *Leishmania.*

Keywords: Trypanosomatids. Gene expression. PABP. RBP

LISTA DE ILUSTRAÇÕES

Revisão de Literatura

Artigo 1

Artigo 2

Artigo 3

Discussão

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

4E-BP	Proteína de ligação ao eIF4E (*4E-Binding protein*)
Alba/ALBA	Proteína/Domínio de acetilação de baixa afinidade de ligação (*Acetylationb lowers binding affinity*)
ATP	Trifosfato de adenosina (*Adenosine triphosphate)*
CAF	Fator associado ao complexo CCR4 (*CCR4-associated fator*)
CCCH	Três resíduos de císteina e um de histidina (**C**ysteine-**C**ysteine-**C**ysteine-**H**istidine)
CCR4	Receptor 4 de quimiocina com motivo C-C (*C-C Motif Chemokine Receptor 4*)
CDS	Sequencia codificadora (*coding sequence*)
DHH	RNA helicase dependente de ATP (*DexD/H-box Helicase*)
DIC	Contraste de interferência diferencial (*Differential interference contrast*)
DNA	Ácido Desoxirribonucléico (*Deoxyribonucleic acid*)
Dox	Doxiciclina (*Doxycycline*)
DRBD	Domínio duplo de ligação ao RNA (*Doubled RNA-Binding Domain*)
dsRNA	RNA de fita dupla (*Double strand RNA*)
DTT	Ditiotreitol (*Dithiothreitol*)
EDTA	Ácido Etilenodiaminotetracético (*Ethylenediaminetetraacetic acid*)
eEF	Fator de alongamento eucariótico (*eukaryotic Elongation Factor*)
eIF	Fator de iniciação eucariótico (*eukaryotic Initiation Factor*)
eRF	Fator de liberação eucariótico (*eukaryotic Release Factor*)
eYFP	Proteína fluorescente amarela aprimorada (*Enhanced Yellow Fluorescent Protein*)
GO	Ontologia gênica (*Gene Ontology*)
GST	Glutationa S-transferase (*Glutathione S-transferase*)
GTP	Trifosfato de guanosina (*Guanosine triphosphate*)
HA	Hemaglutinina (*Hemagglutinin*)
HAT	Tripanossomíase Humana Africana (*Human African Trypanosomiasis*)
HEPES	Ácido hidroxietil piperazinaetanossulfónico (*hydroxyethyl*

	piperazineethanesulfonic acid)
His	Histidina (*Histidine*)
IP	Imunoprecipitação (*Immunoprecipitation*)
IPTG	Isopropil-*β*-D-1-tiogalactopiranosídeo (*Isopropyl-β-D-1-thiogalactopyranoside*)
KH	Domínio de homologia ao RNP nuclear K (*K homology domain*)
LARP	Proteína relaciona ao domínio La (*La-related protein*)
LB	Meio Luria-Bertani (*Luria-Bertani broth*)
LFQ	Quantificação livre de marcação (*Label-Free Quantification*)
MEME	Ferramentas de análise de sequência baseadas em motivos (*Multiple Em for Motif Elicitation*)
Met	Metionina (*Methionine*)
MLLE	Domínio *Mademoiselle*, denominado a partir do motivo conservado KITGMLLE (**M**ethionine-**L**eucine-**L**eucine-glutamat**E**)
mRNA	RNA mensageiro (*Messenger RNA*)
MRP	Proteína de ligação ao RNA mitocondrial (*Mitochondrial RNA binding protein*)
mTORC	Complexo do alvo da rapamicina em mamíferios (*mammalian target of rapamycin complex*)
NOT	Proteína de regulação negative da transcrição (*Negative regulator of transcription protein*)
NTF	Fator de transporte nuclear (*Nuclear transport factor*)
PABC	Domínio C-terminal da PABP (*C-terminal domain of PABP*)
PABP	Proteína de ligação a cauda poli-A (*Poli(a)- Binding Protein*)
Paip	Proteína de interação com a PABP (*PABP-interacting protein*)
PAM2	Motivo 2 de interação a PABP (PABP-interacting motif 2)
PBP1	Proteína de ligação a PAB1 (*Polyadenylate-binding protein*)
PBS	Tampão fosfato-salino (*Phosphate Buffered Saline*)
PCR	Reação em cadeia da polimerase (*Polymerase chain reaction*)
PIC	Complexo de pré-iniciação (*Pre-initiation complex*)
Pol II	RNA polimerase II (*RNA polymerase II*)
PoliPi	Polipirimidinas (*Polypyrimidines*)
PPR	Proteínas de repetição de pentatricopeptídeos (*Pentatricopeptide*

	repeat protein)
PUF	Domínios Pumílio e FBF (Pumilio/*fem-3 binding factor*)
PVDF	Fluoreto de polivinilideno (*Polyvinylidene difluoride*)
qPCR	PCR quantitativa (*PCR quantification*)
RBP	Proteína de ligação a RNA (RNA-Binding protein)
RGG	A**R**ginina-**G**licina-**G**licina (*A**R**rginine-**G**lycine-**G**lycine*)
RNA	Ácido Ribonucleico (*Ribonucleic Acid*)
RNAi	RNA de interferência (*RNA interference*)
RNAseq	Sequenciamento de RNA (*RNA sequencing*)
RNP	Partícula de ribonucleoproteína (*Ribonucleoprotein particle*)
RP	Proteína ribossomal (*Ribosomal protein*)
RRM	Motivo de reconhecimento do RNA (*RNA Recognition Motif*)
SDS-PAGE	Gel de poliacrilamida em condições desnaturantes *(sodium dodecyl sulphate-polyacrylamide gel electrophoresis*)
SFB	Soro Fetal Bovino (*Fetal Bovine Serum*)
SKP	Proteína associada a quinase da fase S (*S-phase kinase-associated protein*)
SL	Sequência Líder (*Splicing Leader*)
SNC	Sistema Nervoso Central (*Central Nervous System*)
SOLiD	Sequenciamento por detecção e ligação de oligonucleotídeos **(***Sequencing by Oligonucleotide Ligation and Detection*)
T7pol	T7 RNA polimerase (*T7 RNA polymerase*)
TetR	Repressor da tetraciclina (*Tetracycline repressor*)
TOP	Sequência de oligopirimidinas na região 5’ terminal (*5’ terminal oligopyrimidines tract*)
tRNAi	RNA transportador iniciador (*Initiation transfer RNA*)
TP-SP	Mutante de fosforilação da PABP1 (*PABP1 Phosphorylation mutant*)
TRRM	Três motivos do tipo RRM (*Three RNA Recognition Motif*)
TSR1	Fator de splicing (*T. brucei serine-arginine (SR) protein*)
UTP	Unidade de transcrição policistrônica (*Polycistronic Transcription Unit*)
UTR	Região não traduzida (Untranslated region)
ZC3H	Domínio dedo de zinco (*Zinc finger CCCH-type*)

SUMÁRIO

1 INTRODUÇÃO

Dentre os protozoários parasitas unicelulares conhecidos como tripanossomatídeos, aqueles pertencentes aos gêneros *Leishmania* e *Trypanosoma* são os mais importantes do ponto de vista médico. As diversas espécies de *Leishmania* são responsáveis pelas diferentes formas de leishmanioses enquanto que *Trypanosoma brucei* causa a doença de sono e *Trypanosoma cruzi*, a doença de Chagas. Essas doenças negligenciadas afetam milhões de indivíduos ao redor do mundo, principalmente em regiões tropicais e subtropicais com baixo desenvolvimento econômico. Os tripanossomatídeos divergiram muito cedo da linhagem evolutiva dos eucariotos e apresentam características genéticas únicas, sobretudo no processo de transcrição caracterizado pela quase total ausência dos promotores da RNA polimerase II. Devido à essa ausência, o controle da expressão gênica acontece basicamente ao nível pós-transcricional, seja pelo controle da estabilidade dos mRNAs e/ou da tradução.

A sobrevivência e adaptação de tripanossomatídeos a novos ambientes requer a ativação de mecanismos de controle pós-transcricionais e de proteínas que se ligam a mRNAs específicos influenciando tanto na sua tradução quanto na degradação. Em ambos os processos, participa a proteína de ligação à cauda poli-A, PABP (*poly-A binding protein*), que é estruturalmente conservada em vários organismos. Em *Leishmania* foram descritos três homólogos distintos (PABP1, PABP2 e PABP3), sendo os dois primeiros também presentes no gênero *Trypanosoma.* A princípio a PABP1 seria a principal candidata a desempenhar funções conhecidas das PABPs de outros organismos durante a tradução. Já o papel funcional das PABP2 e PABP3 ainda é incerto, embora seja conhecido que essas proteínas interagem entre si e se ligam a um conjunto distintos de mRNAs daqueles que se ligam a PABP1.

Diversas outras proteínas de ligação a RNA, conhecidas como RBPs, foram identificadas em tripanossomatídeos e seu estudo tem se tornado cada vez mais importante uma vez que essas proteínas possuem papel chave em diversos processos de controle da expressão gênica, pois se ligam a mRNAs nascentes, maduros e em decaimento. Apesar de sua importância, o conhecimento sobre as RBPs de tripanossomatídeos ainda é escasso. No presente trabalho o principal

objetivo foi caracterizar três dessas proteínas, denominadas RBP23, DRBD2, e ZC3H41, que interagem diferencialmente com as PABPs de *Leishmania infantum*. Como possivelmente essas RBPs participam de complexos diferentes, foram avaliadas a associação ou não com outras proteínas relacionadas ao processo de tradução, e também as interações com seus mRNAs alvos. Os resultados obtidos geraram contribuições importantes quanto ao entendimento de como estas RBPs, e também as PABPs, atuam em processos ligados ao metabolismo dos mRNAs e sua tradução nos tripanossomatídeos.

1.1 Objetivos

1.1.1 Objetivo geral

Caracterizar parceiros proteicos e identificar mRNAs alvos de proteínas de ligação a RNA selecionadas definindo novos mecanismos de controle da síntese proteica em tripanossomatídeos patogênicos.

1.1.2 Objetivos específicos

1. Avaliar os parceiros funcionais da PABP1 de *L. infantum* e, em particular, as RBPs diferencialmente co-precipitadas com os homólogos de PABP;

2. Caracterizar as interações proteicas de duas proteínas de ligação ao RNA, a RBP23 e a DRBD2 e identificar seus mRNAs alvos em *L. infantum*;

3. Contribuir para a caracterização da ZC3H41, uma terceira proteína de ligação ao RNA de *L. infantum*;

4. Investigar a funcionalidade da RBP23 e as proteínas parceiras em *T. brucei* através de análise de fenótipo por RNAi.

5. Propor um modelo para a regulação dos mRNAs associados a RBP23, PABP1 e ZC3H41.

2 REVISÃO DA LITERATURA

2.1 Os tripanossomatídeos

Os tripanossomatídeos são protozoários flagelados pertencentes a classe Kinetoplastida a qual é representada por duas principais ordens: a ordem Bodonida, que compreende organismos de vida livre como o *Bodo saltans*; e a ordem Trypanosomatida, que compreende a família Trypanosomatidae, propriamente. Esta família é formada por nove gêneros distintos (*Blastocrithidia, Crithidia*, *Leishmania, Leptomonas, Paratrypanosoma*, *Phytomonas, Sergeia, Trypanosoma* e *Wallaceina*), diferenciados de acordo com suas ultraestruturas e análises filogenéticas (CAVALIER-SMITH, 2016).

O cinetoplasto, estrutura que deu o nome à classe Kinetoplastida, é uma região especializada da única e grande mitocôndria que comporta DNA mitocondrial bastante complexo e não usual na natureza (CAVALCANTI; DE SOUZA, 2018). Outra estrutura comum à maioria desses organismos é o flagelo único, originado a partir de uma bolsa flagelar na membrana celular localizada próximo da mitocôndria (STUART *et al*., 2008). Morfologicamente, esses organismos podem ser classificados de acordo com várias características como o tamanho e a posição do flagelo, a forma do corpo celular e a distância entre o cinetoplasto e o núcleo. As seis formas morfológicas mais comuns são: amastigota, tripomastigota, epimastigota, promastigota, coanomastigota e opistomastigota (Figura 1). As alterações de uma forma para outra podem ocorrer a depender do ambiente, da forma de cultivo e da adição de drogas (LOPES *et al*., 2010).

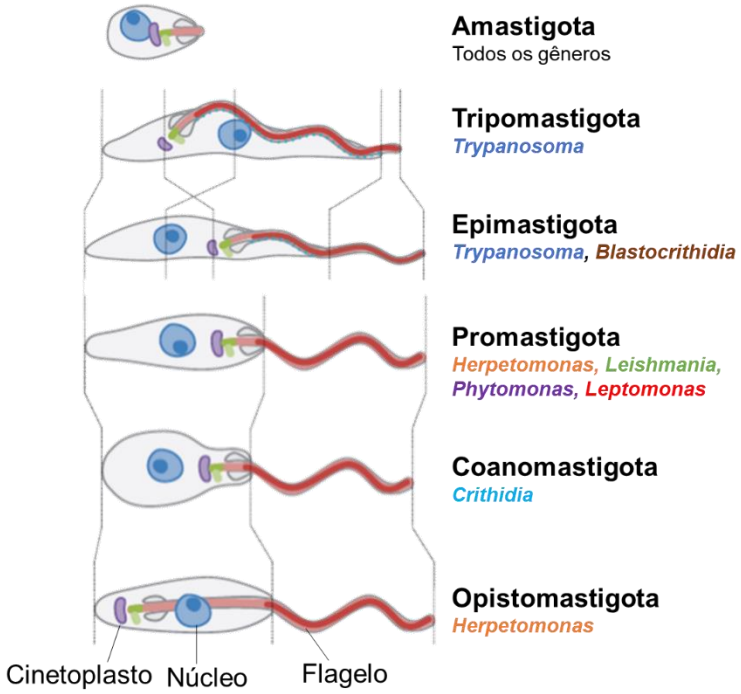


Figura 1. Diagrama das principais classes morfológicas dos tripanossomatídeos. Amastigota, morfologia que não possui um longo e móvel flagelo; Tripomastigota e Epimastigota com um flagelo preso lateralmente ao corpo celular; Promastigota, Coanomastigota e Opistomastigota, morfologia com flagelo livre. Adaptado de WHEELER *et al*., 2013.

Além das variadas características de ultraestrutura celular, os tripanossomatídeos também possuem uma variedade de características moleculares únicas que os diferem dos outros eucariotos e que serão discutidas após uma visão geral das espécies mais relevantes nesse estudo.

2.2 *Leishmania sp*. e as leishmanioses

O gênero *Leishmania* compreende aproximadamente 53 espécies distribuídas em cinco subgêneros: *Leishmania*, *Viannia*, *Sauroleishmania*, complexo *Leishmania enriettii* e *Paraleishmania*. Dessas, 31 espécies são conhecidas por parasitar mamíferos, sendo 20 espécies patógenas aos seres humanos, as quais pertencem principalmente aos gêneros *Leishmania* e *Viannia*. Esses parasitas são organismos heteroxênicos, ou seja, capazes de colonizar dois hospedeiros, vivendo tanto nos fagócitos do retículo endotelial de mamíferos, na forma amastigota, quanto no trato intestinal do inseto vetor, na forma promastigota (AKHOUNDI *et al.,* 2016). Seus

vetores são insetos da ordem Díptera, distribuídos em seis gêneros de acordo com a distribuição geográfica. No Velho Mundo são agrupados os gêneros *Phlebotomus, Sergentomya e Chinius*; e no Novo Mundo (nas Américas) os gêneros *Lutzomyia, Brumptomia* e *Warileya.* Tradicionalmente os gêneros *Phlebotomus* e *Lutzomyia* representam o Velho (Europa, África e Ásia) e o Novo (As américas) mundo, respectivamente (DVORAK *et al*., 2018; BATES, 2007).

Durante o repasto sanguíneo, as formas promastigotas metacíclicas (infectivas) de *Leishmania* presentes no intestino médio do inseto vetor são regurgitadas e inoculadas no hospedeiro mamífero e podem invadir ativamente os macrófagos ou serem fagocitadas (HARHAY *et al*., 2011). Nos macrófagos, os vacúolos fagocitários gerados fundem-se às vesículas lisossomais, formando os fagolisossomos onde os parasitas se diferenciam na forma amastigota. Esta diferenciação é induzida pelas diferenças de temperatura, pH e disponibilidade de aminoácidos e oxigênio entre hospedeiro invertebrado e vertebrado (BESTEIRO *et al*., 2007; MCCONVILLE *et al*., 2007). As formas amastigotas se multiplicam intensamente por divisão binária até o rompimento das células e se disseminam pela corrente sanguínea, permitindo a infecção de novas células e tecidos (STUART *et al*., 2008). Os macrófagos parasitados com as formas amastigotas podem ser ingeridos pelo inseto vetor, durante um novo repasto sanguíneo em mamíferos infectados. No intestino médio do inseto, os amastigotas se diferenciam em promastigotas procíclicas, que após alguns dias de sucessivos ciclos de divisão binária, começam a se replicar mais lentamente e se especializam na forma infectiva, reiniciando o ciclo de transmissão (Figura 2) (KAYE; SCOTT, 2011).

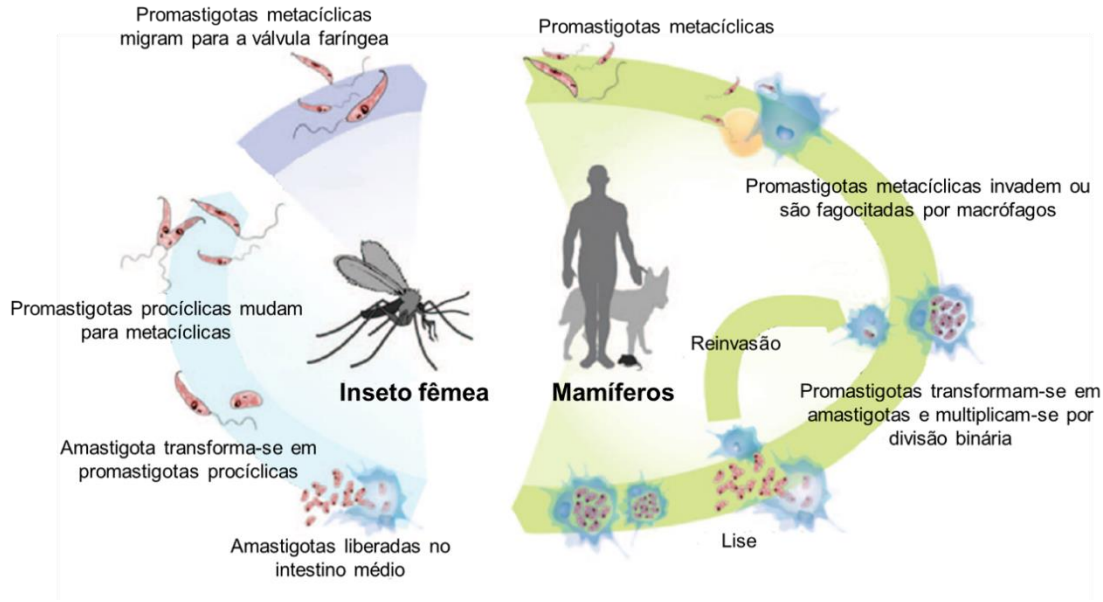


Figura 2. Ciclo de vida de *Leishmania* sp. O inseto infecta o hospedeiro durante a alimentação com as formas promastigotas metacíclicas do parasita. Essas formas invadem ativamente ou são fagocitadas por macrófagos, gerando vacúolos que se fundem com vesículas lisossomais, formandos os fagolisossomos. Nesses ocorre a diferenciação em amastigotas que se multiplicam por divisão binária até ocorrer a lise da célula permitindo os parasitas invadir novos macrófagos. Durante a alimentação, o inseto ingere esses macrófagos com as formas amastigotas que se transformam em formas promastigotas procíclicas no intestino médio, proliferam e se diferenciam nas formas infectantes. Fonte: Adaptado de (KAYE; SCOTT, 2011).

As leishmanioses, doenças causadas pela infecção com diferentes espécies de *Leishmania*, apresentam três principais formas clínicas: cutânea, mucocutânea e visceral. A leishmaniose cutânea é a forma mais comum, causada pelas espécies de dois subgêneros: *Leishmania*, como as *L. major, L. tropica, L. aethiopica*, *L. mexicana* e *L. amazonensis*; e *Viannia*, que inclui as espécies *L. braziliensis*, *L. peruviana*, *L. guyanensis* e *L. panamensis* (Figura 3). Como o próprio o nome diz, essa manifestação clínica da doença é caracterizada pelo aparecimento de lesões na pele, principalmente em partes expostas do corpo. Outra forma da leishmaniose é a mucocutânea, geralmente causada pela *L. braziliensis* e que leva a destruição parcial ou total das membranas mucosas do nariz, boca e garganta. Por último, a leishmaniose visceral, causada principalmente pelas espécies *L. donovani* e *L. infantum*, ambas do subgênero *Leishmania*, é a forma mais severa das três e pode causar crises irregulares de febre, perda de peso, aumento do baço e fígado e anemia, podendo ser fatal se não tratada (WHO, 2018; BATES, 2007).

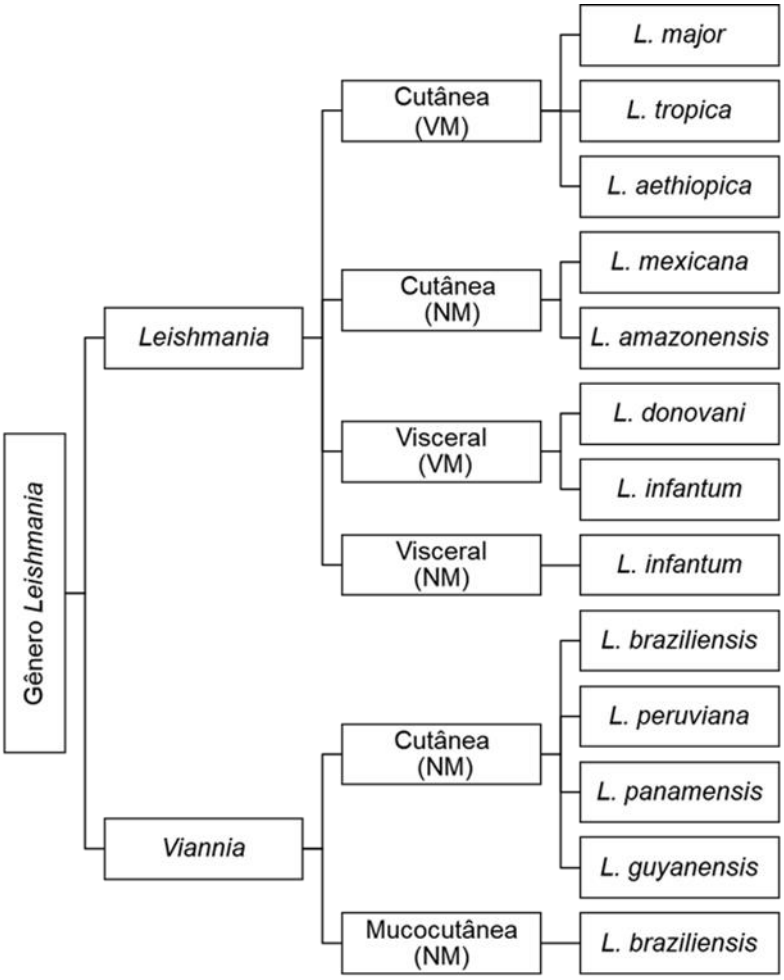


Figura 3. Diagrama das formas da leishmaniose e as principais espécies envolvidas. Principais espécies patogênicas de acordo com a forma da doença que provoca (cutânea, mucocutânea ou visceral) e da distribuição geográfica do vetor no Velho Mundo (VM) para o gênero *Phlebotomus* e Novo Mundo (NM) para *Lutzomyia***.** Adaptado de BATES, 2007.

As leishmanioses são endêmicas de áreas tropicais, subtropicais e da bacia mediterrânea, incluindo mais de 97 países. A maioria dos casos registrados de leishmaniose cutânea ocorrem nos países: Afeganistão, Argélia, Brasil, Colômbia, Irã, Paquistão, Peru, Arábia Saudita e Síria. Em 2017 foram registrados 17.528 casos de leishmaniose cutânea no Brasil, um aumento significativo em relação ao ano anterior no qual foram registrados 12.690 casos. Quase 90% dos casos da leishmaniose mucocutânea ocorrem na Bolívia, Brasil e Peru. Em 2017, mais de 90% dos novos casos registrados de leishmaniose visceral ocorreram em sete países: Brasil, Etiópia, Índia, Quênia, Somália, Sudão e Sul do Sudão. Na região das Américas, 96% dos casos da doença ocorre no Brasil, e em 2017 foram registrados 4.103 casos no país (BRASIL, 2019a; BRASIL, 2019b; WHO, 2018).

O principal tratamento para a leishmaniose visceral ocorre através do uso das drogas anti-Leishmania, juntamente com suporte nutricional, transfusões de sangue e tratamento de outras doenças infecciosas que podem ocorrer simultaneamente. De forma similar, as leishmanioses cutânea e mucocutânea (tegumentar) são tratadas com as drogas anti-Leishmania*,* seja sistemicamente ou topicamente com cremes e pomadas e injeções intralesionais, também sendo usados tratamentos físicos como a crioterapia (BLUM *et al*., 2018; MONGE-MAILLO; LÓPEZ-VÉLEZ, 2018).

O tratamento tradicional com as drogas anti-*Leishmania* começou primeiramente com o uso de compostos antimoniais pentavalentes, porém o desenvolvimento de resistência com taxas maiores de 60% e uma alta toxicidade mostraram a necessidade de pesquisas por novos medicamentos. Outras drogas, mais eficazes e com toxicidade mais baixa, surgiram no mercado, como a anfotericina B, porém seu elevado custo reduz o acesso em países mais pobres. Uma nova opção estudada é a paromomicina com uma toxicidade mediana e de baixo custo (MONGE-MAILLO; LÓPEZ-VÉLEZ, 2018).

2.3 *Trypanosoma brucei* e a doença do sono

Existem muitas espécies no gênero *Trypanosoma,* mas apenas duas se destacam devido à importância médica, que são o *Trypanosoma brucei* e o *Trypanosoma cruzi*, causadores da tripanossomíase africana humana (HAT), também chamada de doença do sono, e da doença de Chagas, respectivamente. Três subespécies são relatadas para *T. brucei*, das quais duas responsáveis pela doença do sono: *Trypanosoma brucei gambiense* e *Trypanosoma brucei rhodesiense*, ambas morfologicamente indistinguíveis. Uma terceira subespécie, *Trypanosoma brucei brucei* não patogênica a seres humanos, parasita animais domésticos e é comumente usada como modelo experimental.

As subespécies de *T. brucei* são transmitidas pela mosca tsé-tsé, do gênero *Glossina*, e atualmente são encontradas quase que exclusivamente no continente africano ao sul do Saara (WHO, 2018; CECCHI *et al*., 2015). No hospedeiro mamífero a forma do *T. brucei* conhecida como sanguínea, mais alongada e delgada, duplica a cada sete horas por fissão binária. Após mecanismos associados

a densidade populacional, essa forma sanguínea alongada e delgada se torna mais curta e ligeiramente compactada, uma pré-adaptação para a vida no intestino médio da mosca tsé-tsé. Uma vez no intestino médio do inseto essa forma se diferencia na forma procíclica que migra para o proventrículo, onde ocorre a diferenciação na forma epimastigota e mais tarde, nas glândulas salivares, em metacíclica. Esta última forma é então capaz de reinfectar o hospedeiro mamífero quando a mosca tsé-tsé faz um novo repasto sanguíneo (Figura 4) (SMITH *et al*., 2017; MATTHEWS, 2009).

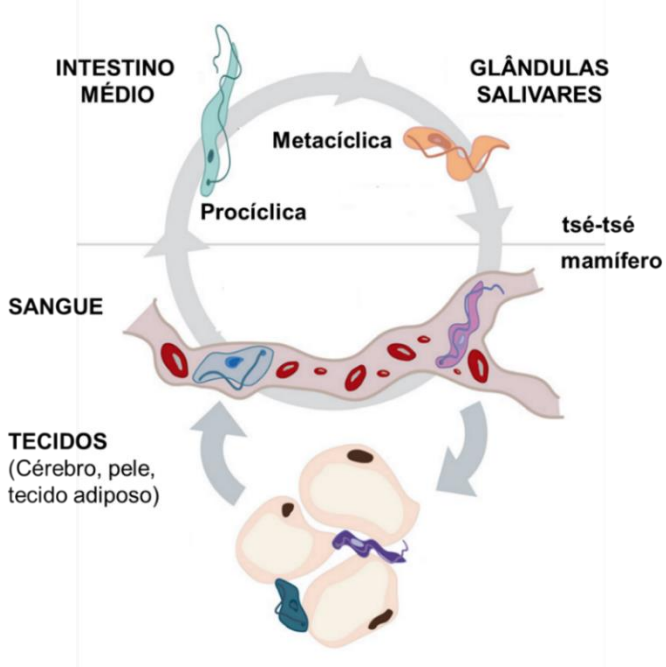


Figura 4. Ciclo de vida do *Trypanosoma brucei*. O ciclo de transmissão abrange dois hospedeiros: um mamífero (humano, bovino e etc.) e a mosca tsé-tsé. No hospedeiro mamífero a forma sanguínea delgada coloniza sangue e tecidos e muda para uma forma mais compactada, a qual é ingerida durante o repasto sanguíneo. No intestino médio do inseto, a forma sanguínea se diferencia em procíclicas que migrarão para as glândulas salivares e se diferenciarão em metacíclicas, a forma infectiva. Adaptado de SMITH *et al*., 2017.

A doença do sono apresenta duas formas clínicas bem distintas: a crônica e a aguda. Na forma crônica o paciente pode ser infectado por meses ou anos de forma assintomática. Quando os sintomas aparecem normalmente é no estágio mais avançado da doença, o qual afeta o sistema nervoso central (SNC). Na forma aguda os sintomas são observados com poucos meses ou até mesmo semanas após a infecção e a doença vai progredindo rapidamente com a invasão do SNC pelo parasita, podendo ser claramente observados os dois estágios da doença. No

primeiro, chamado de hemo-linfático, o parasita se multiplica nos tecidos no sangue e na linfa, e se caracteriza por febres, adenopatias, esplenomegalia, distúrbios no fígado, erupções cutâneas e prurido com muita coceira. No segundo estágio da doença conhecido como meníngeo-encefálico, o parasita atravessa a barreira hematoencefálica para infectar o SNC. Esse estágio está associado a mudanças de comportamento, confusão, problemas na coordenação motora e mental. O distúrbio no ciclo do sono é uma das principais características e é o sintoma que nomeia a doença (PONTE-SUCRE, 2016; BARRETT *et al*., 2003).

A doença do sono está distribuída em 36 países da África Subsaariana, e a sua forma clínica varia dependendo da subespécie envolvida. A forma crônica é mais comum, causada pelo *T. brucei gambiense*, com 97% dos casos registrados em 24 países do oeste e do centro africano. A forma aguda, causada pelo *T. brucei rhodesiense*, é responsável por menos de 3% dos casos em 13 países do leste e sul africano (Figura 5) (WHO, 2018; FRANCO *et al*., 2014). A subespécie *T. brucei brucei,* é um dos agentes causadores da doença que ocorre no gado conhecida como nagana, ou tripanossomíase africana animal, que leva a grandes perdas na pecuária, gerando um grave problema econômico (GIORDANI *et al*., 2016).

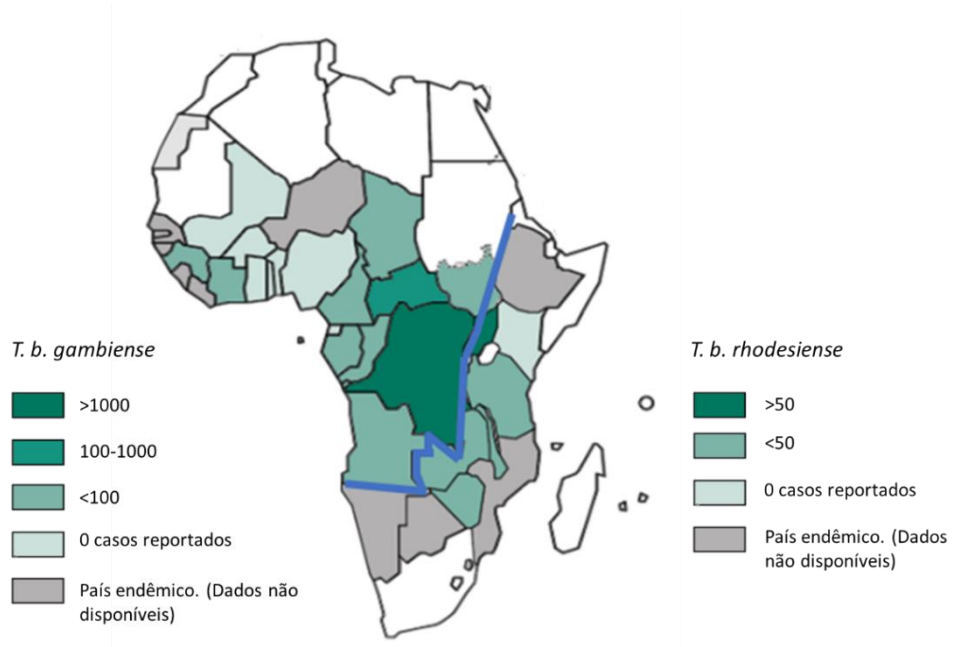


Figura 5. Distribuição da tripanossomíase africana humana em 2014. Número de casos reportados para as duas formas da doença causadas por subespécies de *T. brucei*. A linha azul separa os casos reportados para cada uma das subespécies, a esquerda os países endêmicos para *T. b. gambiense* e a direita os países endêmicos para *T. b rhodesiense*. Os países em branco não são endêmicos para nenhuma das subespécies. Adaptado de WHO, 2018.

O tratamento da doença do sono é bastante difícil devido a elevada toxicidade e a complexa administração das cinco drogas disponíveis: pentamidina, suramina, merlasoprol, eflornitina e nifurtimox. As duas primeiras são usadas nos estágios iniciais da infecção de *T. b. gambiense* e *T. b. rhodesiense*, respectivamente. Já a eflornitina é usada no segundo estágio da infecção de *T. b. gambiense*, porém é ineficaz contra *T. b. rhodesiense*. Equipamentos e pessoal qualificado são necessários para a administração dessa droga, que pode ser usada em combinação com o nifurtimox, o que reduz a duração do tratamento e o número de perfusões. O único tratamento disponível para tratar o segundo estágio de *T. b. rhodesiense* é o merlasoprol que também pode ser usado para *T. b. gambiense*, contudo é a segunda droga indicada para essa subespécie devido a sua alta toxicidade (BAKER; WELBURN, 2018; WHO, 2018).

2.4 Mecanismos moleculares atípicos dos tripanossomatídeos

Os tripanossomatídeos possuem mecanismos de expressão gênica diferenciados dos demais eucariotos. Os genes que codificam proteínas da maioria dos eucariotos, por exemplo, são transcritos em unidades monocistrônicas sob o controle de sequências promotoras específicas. Em tripanossomatídeos, contudo, estes genes são transcritos em longos arranjos policistrônicos, e os transcritos primários são processados por *trans splicing* e poliadenilação que afetam respectivamente as extremidades 5’ e 3’ dos mRNAs (MICHAELI, 2011). Devido a uma escassez de promotores para a RNA polimerase II (pol II), a expressão da maioria dos genes transcritos por essa polimerase é regulada pós-transcricionalmente, seja pelo controle da estabilidade de mRNAs maduros e/ou pelo controle da sua tradução (CLAYTON; SHAPIRA, 2007). Nos tópicos abaixo, esses mecanismos diferenciados serão melhores descritos:

2.4.1 Genoma e transcrição policistrônica

Uma das principais variações entre os genomas dos tripanossomatídeos é em relação a ploidia e ao número de cromossomos. Como exemplo, enquanto que o genoma haploide de *L. major* é composto por 36 cromossomos, o genoma de *T.*

brucei é composto por 11 pares de cromossomos (KAZEMI, 2011; BERRIMAN *et al*., 2005). Contudo, mais de 90% dos genes ortólogos entre esses organismos estão em sintenia, ou seja, se encontram dispostos na mesma organização no genoma (EL-SAYED *et al*., 2005). Além disso, os genomas dos tripanossomatídeos de uma forma geral são caracterizados por uma alta densidade de genes, a presença de longos arranjos gênicos e a quase completa ausência de introns (SIEGEL *et al*., 2010; ULIANA *et al.,* 2008). Estudos comparando os genomas das várias espécies do gênero *Leishmania* e *Trypanosoma* mostram os genes codificadores de proteínas organizados nesses grandes arranjos de genes equivalentes, também conhecidos como unidades de transcrição policistrônicas.

Devido ao fato dos promotores específicos para a RNA polimerase II estarem ausentes ou serem muito raros nos genomas dos tripanossomatídeos, um único sítio de início de transcrição é capaz de conduzir a transcrição de múltiplos genes de forma constitutiva e bidirecional. Um exemplo representativo ocorre no cromossomo 1 de *L. major* onde foram descobertos 79 genes codificadores de proteínas dirigidos por uma única região promotora, sendo as 29 primeiras ORFs (do inglês *open read frame*) todas codificadas em uma das fitas e as 50 ORFs restantes na fita oposta (TEIXEIRA *et al*., 2012; MYLER *et al*., 1999).

Apesar de uma aparente similaridade com os operons bacterianos, a grande diferença nos tripanossomatídeos é que os genes co-transcritos nas unidades policistrônicas não codificam proteínas funcionalmente relacionadas. Assim, em contraste com a maioria dos organismos, os tripanossomatídeos parecem ter perdido ou nunca adquirido a capacidade de regular a iniciação da transcrição de genes individuais (TEIXEIRA *et al*., 2012; REQUENA, 2011; MARTÍNEZ-CALVILLO *et al*., 2010).

2.4.2 Formação de mRNAs maduros

Após a produção do pré-mRNA policistrônico de tripanossomatídeos, esse é processado para a formação de mRNAs maduros individuais. Na maioria dos eucariotos o processamento dos mRNAs ocorre em três etapas: (1) adição da estrutura *cap* (7-metil guanosina) na extremidade 5’ do mRNA, logo após o início da transcrição; (2) remoção das sequências não codificantes, também chamadas de

introns, pelo mecanismo de *cis-splicing*; (3) adição de uma sequência de 100-200 adeninas na extremidade 3', após um sinal de poliadenilação que em mamíferos consiste na sequência AAUAAA (SNUSTAD; SIMMONS, 2013). Em tripanossomatídeos, os mRNAs precursores de arranjos policistrônicos são processados pelos mecanismos de *trans-splicing*. Estes ocorrem de forma simultânea e acoplados a poliadenilação (Figura 6), sendo ambos dirigidos por sequências comuns ricas em polipirimidinas presentes nas regiões intergênicas (BENZ *et al.,* 2005, LIANG *et al*., 2003). No *trans-splicing,* ocorre a adição na extremidade 5' do mRNA de um mini-éxon contendo 39 nucleotídeos, chamado de sequência líder (SL). Essa sequência é caracterizada pela presença do *cap* modificado pela metilação dos quatro primeiros nucleotídeos, assim denominado *cap4*. Curiosamente a RNA pol II reconhece o promotor da sequência líder e seus precursores, de maior tamanho, são transcritos individualmente (CRIBB; SERRA, 2009).

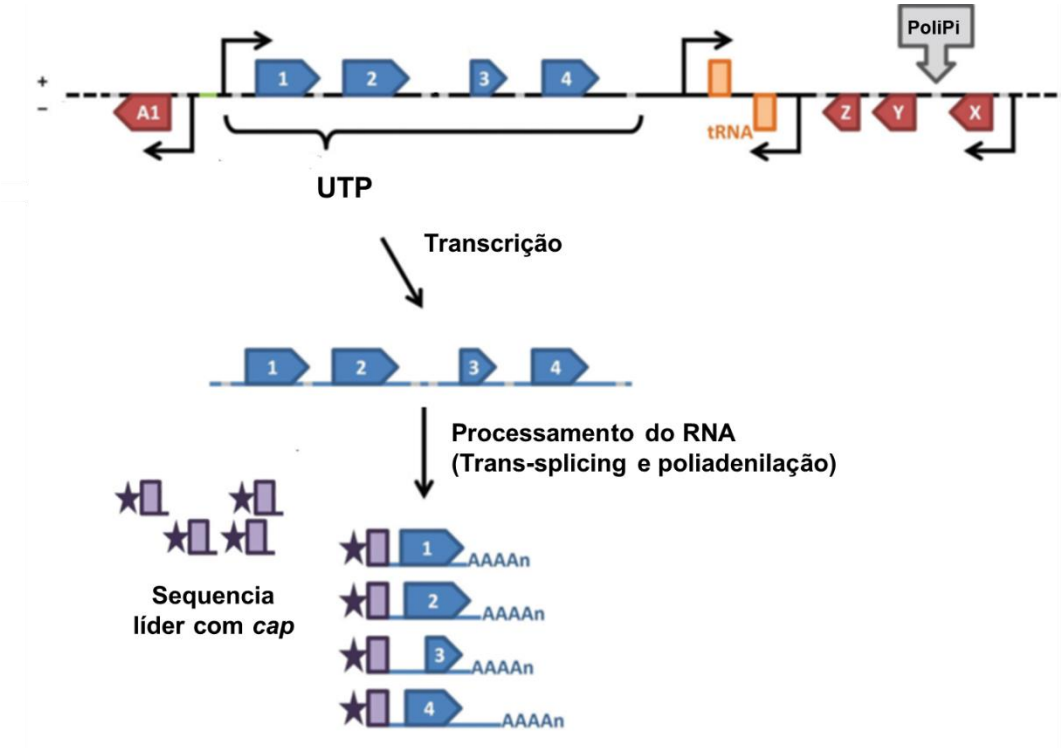


Figura 6. Transcrição gênica e processamento dos mRNAs em tripanossomatídeos. Grandes agrupamentos de genes não relacionados (setas azuis) estão organizados em unidades de transcrição policistrônicas (UTP). Os RNAs policistrônicos (pré-mRNAs) são individualizados em mRNAs monocistrônicos após a adição da sequência líder com *cap* (caixa e estrela roxa) através do *trans-splicing* acoplado a poliadenilação. Esse processamento é guiado por uma sequência de polipirimidinas (PoliPi) que estão presentes na região intergênica. Fonte: Adaptado de TEIXEIRA *et al*., 2012.

No processo de poliadenilação nos tripanossomatídeos ocorre a adição na extremidade 3’ do mRNA de uma sequência consecutiva de adeninas, formando a cauda poli-A, a uma distância fixa (100 a 400 nucleotídeos) após o sinal de processamento. Os dois processos, *trans-splicing* e poliadenilação, permitem então que o mRNA policistrônico precursor gere mRNAs monocistrônicos maduros (HAILE; PAPADOPOULOU, 2007; MARTÍNEZ-CALVILLO *et al.,* 2010; MICHAELI, 2011). Eventos de *trans-splicing* alternativos também podem ocorrer nesses organismos, podendo ocasionar mudanças no sitio de adição da sequência líder, nos elementos regulatórios da 5’UTR ou no códon iniciador AUG (NILSSON *et al*., 2010).

2.4.3 Controle da expressão gênica

Em tripanossomatídeos, a falta de controle da transcrição aumenta consideravelmente a importância de processos pós-transcricionais na regulação da expressão gênica. Esses processos agiriam tanto pelas etapas de maturação dos mRNAs, vistas no tópico acima, como também a nível de controle de sua estabilidade e/ou tradução em proteínas (CLAYTON; SHAPIRA, 2007). Após o processamento, o mRNA maduro precisa ser exportado por um complexo ribonucleoproteico (RNP), do núcleo para o citoplasma, para ser traduzido. Ao chegar no citoplasma, o mRNA deve se associar com proteínas requeridas para o transporte citoplasmático para um local subcelular especifico, tradução, armazenamento ou degradação (Requena, 2011). Essas proteínas de ligação ao RNA, conhecidas como RBPs (*RNA-binding proteins*), são elementos de ação *trans* que interagem com motivos regulatórios dos transcritos maduros, os elementos de ação *cis*. Essas proteínas participam de todos os estágios do controle da expressão gênica dos tripanossomatídeos como processamento e estabilidade dos mRNAs e tradução (ERBEN *et al.,* 2014; MÜLLER-MCNICOLL; NEUGEBAUER, 2013).

2.5 Proteínas de ligação ao RNA (RBPs)

Em tripanossomatídeos, como mencionado o controle da expressão gênica é em parte determinado pela ação das proteínas de ligação ao RNA, que podem

influenciar tanto na estabilidade dos transcritos quanto na eficiência da tradução. Essas proteínas irão controlar a vida do mRNA, pois se ligam a RNAs recém sintetizados, maduros e em declínio, formando complexos de ribonucleoproteínas (GLISOVIC *et al.,* 2008).

Diferentes transcritos que contêm os mesmos motivos regulatórios parecem ser regulados de forma similar (NAJAFABADI *et al.,* 2013). Esses motivos são alvos de diferentes RBPs que se associam com os mRNAs formando os complexos de mRNPs. As RBPs interagem com motivos dos mRNAs usando domínios funcionais como RRM (*RNA Recognition Motif*), dedo de zinco (*Zinc Finger*), Pumilio e ALBA, os quais são os mais relevantes domínios encontrados nas RBPs de tripanossomatídeos (KOLEV *et al.,* 2014; CLAYTON, 2013). As RBPs ainda podem ser denominadas DRBDs e TRRMs quando possuem dois e três RRMs, respectivamente. Análises de similaridade de sequência e a comparação da posição dos genes dos genomas dos três principais tripanossomatídeos, conhecidos como TriTryps, identificaram 139 proteínas contendo RRMs em *T. cruzi*, 75 em *T. brucei* e 80 em *L. major*, sendo 77 proteínas ortólogas nos três organismos (GAUDENZI *et al*., 2005).

A regulação para as diferentes RBPs é exercida em relação a diferentes aspectos tais como a expressão em um determinado estágio de desenvolvimento, localização, biogênese dos complexos de RNPs, e condensação das RBPs e mRNAs em grânulos citoplasmáticos. Modificações pós-traducionais e outras proteínas relacionadas ao seu metabolismo podem também modular a abundância dos mRNAs (ROMANIUK *et al.,* 2016).

2.5.1 Proteína de ligação à cauda poli-A (PABP)

A proteína de ligação ao RNA, de maior relevância em todos os eucariotos é a proteína de ligação à cauda poli-A, ou PABP (do inglês *Poly-A Binding Protein*). Essa proteína participa de praticamente todos os eventos mRNA-dependentes devido a sua alta afinidade pelas sequências de adenosina presentes na extremidade 3’ dos mRNAs. Além dos mRNAs, as PABPs interagem com outras proteínas, promovendo assim o seu envolvimento em vários eventos celulares. No núcleo, as PABPs desempenham um papel na poliadenilação, determinando o

comprimento da cauda poli-A e podem estar envolvidas na exportação de mRNAs. No citoplasma, participam na iniciação da tradução além de agirem evitando a degradação dos mRNAs, através da ligação a sua cauda poli-A, ou estimulando seu decaimento, através de interações do mRNA com complexos proteicos com função de deadenilase (GALLIE, 2014; ELISEEVA *et al.,* 2013; GOSS; KLEIMAN, 2013).

A estrutura das PABPs citoplasmáticas típicas caracterizadas até o momento é bastante conservada entre os diversos eucariotos e pode ser dividida em: uma região N-terminal, um segmento de ligação não conservado e uma região C-terminal caracterizada pela presença de um domínio conservado conhecido como MLLE, anteriormente denominado PABC (Figura 7) (ELISEEVA *et al.,* 2013; GOSS; KLEIMAN, 2013; KÜHN; WAHLE, 2004). A região N-terminal das PABPs compreende cerca de dois terços da molécula e é constituída por quatro domínios de ligação ao RNA conservados (RRMs 1, 2, 3 e 4) posicionados em *tandem*. Os RRMs são caracterizados pela presença de dois motivos altamente conservados conhecidos como RNP1 (octâmero) e RNP2 (hexâmero) dentro de uma região estruturalmente, mas não sequencialmente conservada de aproximadamente 80 aminoácidos. O RRM é capaz de formar uma estrutura terciária composta por quatros folhas beta e duas alfa hélices ($\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$) (LOERCH; KIELKOPF, 2016; POMERANZ-KRUMMEL; NAGAI, 2001). Os RRMs 1 e 2 são responsáveis pela ligação específica ao poli-A e também interagem com o fator eIF4G (CHENG; GALLIE, 2007; IMATAKA *et al.,* 1998; KUHN *et al.,* 1996). Os RRMs 3 e 4 ligam a sequências AU não poliméricas, com significado funcional desconhecido e também interagem com outras proteínas como por exemplo, o fator de elongação 1α (eEF1α) (KHACHO *et al*., 2008; KHANAM *et al*., 2006; SLADIC *et al*., 2004). Em mamíferos o segmento de ligação, rico em prolina, glutamina e metionina, parece estar envolvido no processo de multimerização de várias moléculas de PABP na cauda poli-A (LIN *et al.*, 2012; MELO *et al.*, 2003;). Na região C-terminal, o MLLE é um domínio de interação proteína-proteína, que permite a interação com diversas outras proteínas, como por exemplo o fator de terminação eRF3 e as proteínas regulatórias Paip1 e Paip2 de mamíferos (KOZLOV; GEHRING, 2010; MELO *et al.*, 2003; KOZLOV *et al.,* 2001).

Figura 7. Estrutura da PABP. A proteína contém quatro RRMs, uma região de ligação e o domínio MLLE, anteriormente denominado de PABC. Em humanos, os RRMs 1 e 2 (em azul) interagem com o eIF4G e a região do MLLE (ou PABC, em vermelho) interage com o fator eRF3. Fonte: Adaptado de GALLIE, 2014.

Em *Leishmania*, foram encontrados e caracterizados três homólogos da PABP (PABP1, 2 e 3) (DA COSTA LIMA *et al*., 2010; BATES *et al*., 2000). Desses, apenas dois (PABP1 e 2) estão conservados na maioria dos outros tripanossomatídeos uma vez que a PABP3 foi perdida nas espécies de *Trypanosoma* (KRAMER *et al*., 2013). Na presença de inibidores transcricionais as PABP2 e 3 migram para o núcleo enquanto a PABP1 permanece predominantemente no citoplasma, sugerindo uma função nuclear que pode estar associada ao processamento de mRNAs e/ ou seu transporte do núcleo (KRAMER *et al*., 2013; DA COSTA LIMA *et al*., 2010). Apesar das três PABPs terem alta afinidade por RNA, a PABP2 possui substituições de resíduos de aminoácidos críticos para o reconhecimento da poli-A, sugerindo uma especificidade diferencial no reconhecimento de mRNAs alvos (GUERRA *et al*., 2011; DA COSTA LIMA *et al*., 2010; PITULA *et al*., 1998). Do ponto de vista funcional, os estudos dos homólogos de PABP em *Leishmania* indicam que a PABP1 interage com o complexo eIF4F mais diretamente implicado na tradução dos mRNAs nos tripanossomatídeos (DA COSTA LIMA *et al*., 2010), descrito em mais detalhes a seguir.

2.5.2 RBPs e a regulação pós-transcricional

Uma das formas de se controlar a abundância de mRNAs é através da regulação da velocidade de vias de degradação. Em tripanossomatídeos o modelo para degradação de mRNAs envolve dois mecanismos distintos (CLAYTON; SHAPIRA, 2007). O primeiro é rápido, independente de eventos de deadenilação e parece ocorrer em mRNAs instáveis estágio-específicos, e sua regulação envolve mecanismos parasito-específicos. Já o segundo mecanismo é constitutivo, dependente de deadenilação, e necessita de baixa energia para a degradação de

mRNAs estáveis e provavelmente de uma subpopulação de mRNAs instáveis (HAILE; PAPADOPOULOU, 2007). Acredita-se que RBPs específicas ligam-se a motivos regulatórios dos mRNAs e modulam a interação com elementos da maquinaria de degradação (HASAN *et al*., 2014).

Uma outra forma de controlar a abundância dos mRNAs, é através da condensação de distintos complexos RNPs. Em *T. brucei* e *T. cruzi*, sob condições de falta de nutrientes, é formada uma grande quantidade de grânulos citoplasmáticos, formada pela condensação de diferentes complexos RNP. Esses grânulos dificultam a degradação dos mRNAs, e os transcritos armazenados permanecem intactos desde que possuam cauda poli-A e o característico mini-éxon na extremidade 5'. Os transcritos presentes nestes grânulos poderão ser traduzidos uma vez que uma nova fonte de nutrientes esteja disponível (ROMANIUK *et al.,* 2016; CASSOLA *et al.,* 2007).

A tradução é o último passo no fluxo da informação genética e a sua regulação permite uma resposta imediata e rápida às mudanças nas condições fisiológicas. Após a exportação do núcleo para o citoplasma, nem todos os mRNAs entram imediatamente no *pool* ativo de tradução. Muitos são mantidos, em vez disso, em um estado quiescente de tradução, aguardando uma localização subcelular adequada ou algum sinal alertando o momento correto para produzir proteínas (GEBAUER; HENTZE, 2004). A regulação da tradução desempenha um papel crítico durante o desenvolvimento e, em geral, também será mediada por sinais de ação *cis* nos mRNAs alvos e *trans* das RBPs (KUERSTEN; GOODWIN, 2003), atuando principalmente durante a fase inicial do processo.

2.6 Iniciação da tradução

Em eucariotos, das quatro etapas da tradução, iniciação, alongamento, terminação e reciclagem dos ribossomos, a primeira é a principal etapa de regulação. Na iniciação ocorrem eventos importantes como a montagem do ribossomo, reconhecimento do mRNA e identificação do códon AUG de iniciação da tradução (JACKSON, 2010). Nessa etapa estão envolvidas diversas proteínas auxiliares denominadas de eIFs (do inglês, *eukaryotic Initiation Factor*). Primeiro ocorre a formação do complexo ternário entre o eIF2, uma molécula de GTP e um

metionil-tRNA iniciador (eIF2-GTP-Met-tRNAi). Este complexo interage com a subunidade ribossomal 40S através dos fatores eIF1, eIF1A e eIF3, formando o complexo de pré-iniciação 43S. Neste complexo o fator eIF3 é capaz de interagir com o fator eIF4F ligado ao mRNA, formando o complexo de iniciação 48S (HINNEBUSCH, 2014; PRÉVÔT *et al*., 2003).

O fator eIF4F (4F) exerce função essencial na ligação entre o mRNA e o ribossomo, sendo composto por três polipeptídios: (1) eIF4A, uma ATPase RNA-dependente e RNA helicase que remove estruturas secundárias do mRNA, permitindo uma melhor ligação do ribossomo e o rastreamento ao códon AUG de iniciação da tradução; (2) eIF4E, uma proteína que se liga ao *cap* presente na extremidade 5' do mRNA e (3) eIF4G, proteína que possui sítios de ligação para eIF4E, eIF4A, eIF3 e PABP. A ligação entre o eIF4F e a PABP permite a circularização do mRNA, o que parece estimular a tradução e a reciclagem dos ribossomos junto a um mesmo mRNA (Figura 8) (MERRICK, 2015). A regulação da iniciação da tradução pode ocorrer através do controle do acesso do eIF4E a estrutura *cap*, da interação eIF4G-eIF4E, do recrutamento do complexo de iniciação 43S, da montagem do 80S, da elongação e da cauda poli-A. Todas estas etapas podem ser influenciadas pela ação de proteínas ligadas a elementos *cis* na 3'UTR dos mRNAs (TAKEUCHI; YAMASHITA, 2017).

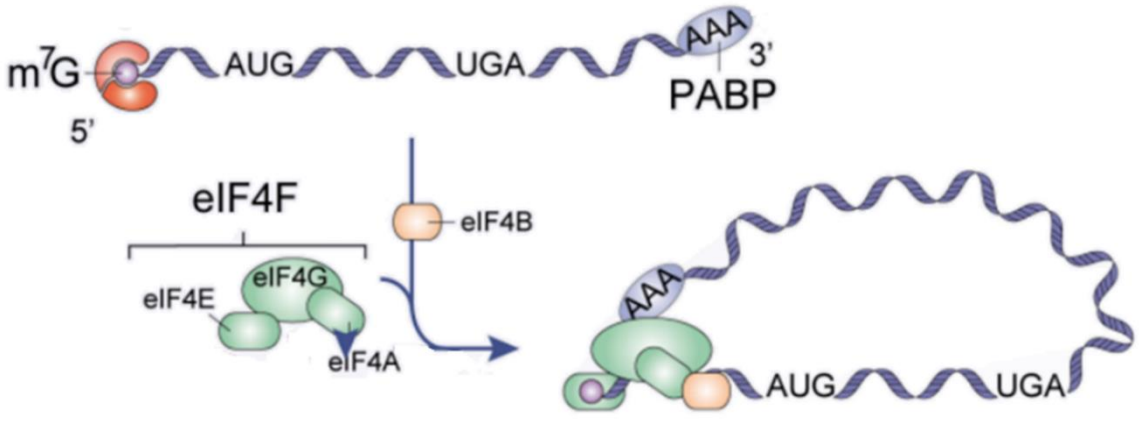


Figura 8. Complexo de iniciação da tradução. A proteína eIF4E, pertencente ao complexo eIF4F, juntamente com as proteínas eIF4A e EIF4G liga-se ao *cap* do mRNA. A PABP ligada a cauda poli-A do mRNA interage com o fator eIF4G e dessa forma permite a formação da conformação circular fechada. Fonte: Adaptado de GUERRERO *et al*., 2015.

Em tripanossomatídeos, diversas proteínas homólogas foram encontradas para a maioria dos fatores envolvidos na iniciação da tradução e descritos nos demais eucariotos. Chama a atenção, entretanto, que dentre os componentes do fator eIF4F foram identificados e caracterizados em *Leishmania e Trypanosoma* dois homólogos para o eIF4A (EIF4AI e EIF4AIII) (DHALIA *et al.,* 2006; DHALIA *et al.,* 2005), seis para o eIF4E (EIF4E1 a 6) (FREIRE *et al.,* 2014; YOFFE *et al.,* 2006; DHALIA *et al.,* 2005; YOFFE *et al.,* 2004;) e cinco para o eIF4G (EIF4G1 a 5) (YOFFE *et al.,* 2009; YOFFE *et al.,* 2006; DHALIA *et al.,* 2005). Ensaios de interação mostraram que dois homólogos do eIF4G (EIF4G3 e EIF4G4) são capazes de interagir com o EIF4AI, enquanto que o EIF4G4 também interage com o EIF4E3 ao passo que o EIF4G3 interage com EIF4E4 (DHALIA *et al.,* 2005; FREIRE *et al.,* 2011). O complexo formado pela interação entre EIF4E4/EIF4G3/EIF4AI teria papel principal na tradução da maioria dos mRNAs, enquanto o complexo formado por EIF4E3/EIF4G4/EIF4AI seria mais seletivo, traduzindo mRNAs específicos (DE MELO NETO *et al.,* 2015; MOURA *et al.,* 2015). A PABP1 interage com o complexo principal EIF4E4/EIF4G3/EIF4AI enquanto que as PABP2 e 3 poderiam interagir com o segundo complexo EIF4E3/EIF4G4/EIF4AI (Figura 9) e também exercer outras funções no núcleo. Uma questão importante para ser investigada é quais as RBPs que irão modular e direcionar os mRNAs que serão traduzidos por um ou outro complexo.

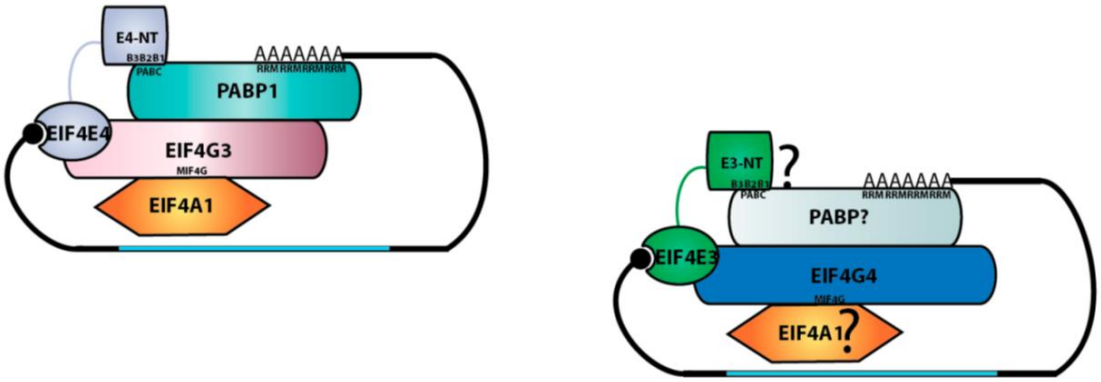


Figura 9. Complexos de tradução do tipo eIF4F dos tripanossomatídeos. O EIF4E4 forma o principal complexo envolvido na tradução em tripanossomatídeos, através da interação com o EIF4G3 (que também interage com EIF4AI) e com a PABP1 por meio de sua porção N-terminal (E4-NT). Uma composição similar é possível para o segundo complexo eIF4F formado por EIF4E3, EIF4G4, EIF4AI e um dos diferentes homólogos da PABP. Adaptado de FREIRE *et al*., 2017.

2.7 Contexto atual do estudo

Quando do início da execução desta tese, as informações mais claras sobre as PABPs de tripanossomatídeos indicavam que essas participariam de complexos distintos, um formado pela PABP1 e outro pela PABP2 e/ou 3, essa última quando presente (Figura 9). Um primeiro grande questionamento era quanto aos conjuntos de mRNAs que estariam associados a cada complexo, se eram iguais ou diferentes. Análises iniciais de RNAseq a partir da imunoprecipitação da proteína nativa com anticorpo policlonal em *L. major* demostraram que a PABP1 estava associada em sua maioria a mRNAs de proteínas ribossomais enquanto as PABP2/ PABP3 estavam enriquecidas com mRNAs de histonas (comunicação pessoal – Fabíola Holetz). As próximas perguntas buscaram a razão e como ocorre essa seleção específica de diferentes conjuntos de mRNAs, além de identificar quais outras proteínas estão envolvidas e suas funções. Com base nisso, estudos do grupo identificaram proteínas hipotéticas contendo domínios de ligação ao RNA, do tipo RRM, diferencialmente associadas a PABP1 e as PABP2 e PABP3 de *L. infantum* (DA COSTA LIMA, 2012 – tese de doutorado). Duas dessas proteínas hipotéticas, codificadas pelos genes *Linj.17.0610* e *Linj.35.2240,* a primeira associada a PABP1 e a segunda a PABP2, foram selecionadas e expressas em *L. infantum* para uma melhor caracterização (SILVA, 2014 – dissertação de mestrado). Em *T. brucei* essas proteínas foram denominadas RBP23 e DRBD2, respectivamente, e apesar de evidências que nesses organismos regulam a expressão do mRNA repórter (LUEONG *et al*., 2016), os mecanismos envolvidos ainda permanecem desconhecidos. Posteriormente mais uma proteína foi selecionada para caracterização, a ZC3H41, que co-precipitada tanto com PABP1 quanto com a RBP23, aumentando assim o painel de proteínas que possam estar envolvidas nesse processo. Na presente tese as três RBPs descritas de *L. infantum*, junto com suas parceiras PABPs, foram selecionadas para uma melhor caracterização e compreensão quanto a eventuais funções mediando ou auxiliando aspectos do metabolismo dos mRNAs, como sua seleção para a tradução. De fato, com a coleção de resultados obtidos, foi proposto um modelo para a seleção de um conjunto específico de mRNAs para a tradução. Ao mesmo tempo, uma vez definido um complexo de proteínas associadas a própria RBP23, buscou-se aplicar algumas

outras técnicas que são deficientes em *L. infantum,* em um segundo modelo experimental, o *T. brucei* para definir melhor sua função. Os dados obtidos em cada capítulo apresentado geraram novos conhecimentos a respeito da função destas RBPs e dos homólogos de PABP bem como seu papel associadas a iniciação da tradução, como um todo, nos tripanossomatídeos.

3 ARTIGO 1

3.1 Identificação de proteínas e mRNAs diferencialmente ligados aos homólogos da proteína de ligação à cauda poli-A de Leishmania revela uma nova associação entre PABP1, RBP23 e mRNAs de proteínas ribossomais.

Artigo anexo na página seguinte.

A doutoranda e primeira autora desse artigo em finalização para submissão de acordo com as normas da revista *Nucleic Acid Research* (encontradas em anexo), foi responsável pela escrita do artigo e pelos experimentos e dados gerados referentes as proteínas RBP23 e a DRBD2. Resumidamente, os dados de espectrometria de massas confirmaram a interação da RBP23 com a PABP1 e o complexo de tradução EIF4E4/EIF4G3 bem como com proteínas hipotéticas específicas. Em contraste, a DRBD2 interage com a PABP2/PABP3 e muitas outras proteínas de ligação ao RNA. A ligação da RBP23 e PABP1 é direta e a região de interação foi mapeada por ensaios de *pull down*. Nas análises de RNAseq, mais de 80% dos mRNAs ligados a RBP23 foram de proteínas ribossomais, os quais estão ausentes nos dados obtidos para a DRBD2.

Identification of proteins and mRNAs differentially bound to the *Leishmania* Poly(A) Binding Proteins reveals a novel and specific association between PABP1, RBP23 and mRNAs encoding ribosomal protein

Ludmila A. Assis[1,2], Tamara D. C. da Costa Lima[3], João R. da Cruz Silva[1], Maria J. R. Bezerra[1], Irassandra R. P. U. C. de Aquino[1], Moezio V. C. Santos Filho[1,2], Camila C. Xavier[1], Kleison C. Merlo[1], Fabiola B. Holetz[4], Christian Probst[4], Antônio M. Rezende[1], Barbara Papadopoulou[5] and Osvaldo P. de Melo Neto*[1]

[1] Department of Microbiology, Aggeu Magalhães Institute – Oswaldo Cruz Foundation, Recife, PE, Brazil.

[2] Federal University of Pernambuco, Recife, PE, Brazil.

[3] University Center Tabosa de Almeida, Caruaru, PE, Brazil.

[4] Carlos Chagas Institute, Curitiba, PR, Brazil.

[5] CHU de Quebec Research Center and Department of Microbiology-Infectious Disease and Immunology, Laval University, Quebec, QC, Canada.

*Corresponding author e-mail: opmn@cpqam.fiocruz.br

Keywords: Trypanosomatidae. RBP23. PABP1.

ABSTRACT

Regulation of gene expression in *Leishmania* depends on the control of mRNA stability and translation and these events require the participation of RNA binding proteins (RBPs). Poly(A) Binding Protein (PABP), a *major* eukaryotic RBP, is known to have multiple roles associated with mRNA and its translation. In *Leishmania*, three PABPs have been identified, with PABP1 found to associate with undefined mRNA targets distinct from PABPs 2 and 3. Here, through co-precipitation and RNA-seq experiments, we found ribosomal protein (RP) mRNAs preferentially bound to PABP1, while PABP2 and PABP3 were enriched with unrelated messengers**.** Co-precipitation and mass-spectrometry analysis also revealed two RBPs differentially associated with PABP1 or PABP2, respectively RBP23 and DRBD2. Reciprocal experiments confirmed PABP1, plus the EIF4E4/EIF4G3 translation complex and specific hypothetical proteins, co-precipitating with RBP23. In contrast, PABP2/PABP3 and many other proteins associated with DRBD2. Both RBPs bound directly to the three PABPs and mapping this interaction on PABP1 found overlapping but non identical binding regions within the PABP1 RRMs. Over 80% of the mRNAs bound to RBP23 were RP-mRNAs, mainly lacking from the DRBD2 bound messages. It instead co-precipitated with mRNAs more consistent with PABP2. These experiments raise the possibility of RBP23 mediating the recruitment of RP-mRNAs by PABP1 and channeling them to translation by the EIF4E4/EIF4G3 complex.

INTRODUCTION

The trypanosomatids protozoans constitute a group of parasitic microorganisms which include several species pathogenic to human, all belonging to the *Leishmania* and *Trypanosoma* genera. These are early divergent eukaryotes characterized by a number of unique features related to their gene expression and mRNA metabolism (1). Noteworthy, their protein coding sequences are transcribed as part of long precursor polycistronic arrays that are subsequently processed by *trans*-splicing and polyadenylation into mature monocistronic mRNAs (2)**.** Unlike other eukaryotes, and due to a lack of promoters for RNA polymerase II, which leads

to the mRNAs being transcribed constitutively, expression of most trypanosomatids genes is regulated post-transcriptionally. It is assumed then that this regulation mainly targets mechanisms associated with the processing, transport, stability and translation of mature mRNAs (3, 4). The trypanosomatids then emerge as relevant models for the understanding of mechanisms mediating such regulatory events in eukaryotes.

RNA binding proteins (RBPs) recognize and bind regulatory sequences in targets RNAs and are required for events such as mRNA processing, subcellular transport, translation, storage and degradation (5–7). These proteins may interact with related sets of RNAs, suggesting an extensive level of regulation (8, 9). Trypanosomatids have an unusual large number of RBPs belonging to distinct functional families and with different mRNA binding domains. These include those with the RNA Recognition Motif (RRM), Zinc Finger, Pumilio and ALBA domains (10–12). Many reports have shown RBPs acting as post-transcriptional regulators of gene expression in trypanosomatids. The RBPs usually bind to sequence elements located within the 3' untranslated regions (3' UTRs) of mature transcripts. Different transcripts containing the same motifs and coding for functionally related proteins appear to be similarly regulated (13).

The cytoplasmic poly(A) binding protein, or simply PABP (distinct from the NPABP, nuclear and more functionally restricted), is the best known and the most abundant of the eukaryotic RBPs. It participates in most, if not all, mRNA associated events, a consequence of its high affinity to the 3' end adenosine tract, the poly(A) tail, found in most eukaryotic mRNAs. PABPs interact with many other proteins, thus promoting their involvement in various cellular events. Despite being mainly cytoplasmic, they are also found within the nucleus, where they may play a role in polyadenylation, determining the length of the poly(A) tail, and may be involved in mRNA export (14, 15). In the cytoplasm, PABP bound to poly(A) tail interacts with the eIF4F translation initiation complex, through its eIF4G subunit, bound to the 5' end of the mRNA, contributing to formation of a closed loop structure. Its interaction with the termination factor eRF3, may also bring together the initiation and stop codons, possibly facilitating translation initiation and/or ribosome recycling (16, 17). In addition, PABP alternatively prevents the degradation of mRNAs, by binding to their

poly(A) tail, or stimulate the mRNA decay through interactions with deadenylase complexes (14, 18).

The primary structure of the PABPs, conserved throughout the known eukaryotes, can be divided into three distinct segments. The N-terminal region comprises roughly two-thirds of the protein and consists of four highly conserved RRM domains (1, 2, 3 and 4). This is followed by a non-structured and variable linker segment, rich in proline, glutamine and methionine in mammals, which has been implicated in the multimerization process seen when several PABP molecules bind to the poly(A) tail (19, 20). The C-terminal segment is mainly characterized by the presence of the conserved C-terminal domain, known as MLLE but previously named PABC, shown to mediate protein-protein interactions (14, 15, 21).

Three PABP homologues were identified in *Leishmania* and other trypanosomatids species (PABP1, PABP2 and PABP3), with PABP3 found to be lost from the *Trypanosoma* lineage (22, 23). *Leishmania* PABP1 co-precipitates with the eIF4F initiation complex found to be most likely involved with the translation of most mRNAs, based on the EIF4E4/EIF4G3 subunits (22, 24). A novel and direct interaction between PABP1 and EIF4E4 has been identified, mediated by three conserved PAM2 motifs within the N-terminus of EIF4E4 that bind to the PABP1 MLLE domain (25). *Leishmania* PABP2 and PABP3 co-precipitate together and after transcription inhibition both proteins migrate to the nucleus, while PABP1 remains predominantly in the cytoplasm (22), similar in *T. brucei* in which only PABP2 but not PABP1 can accumulate in the nucleus (23). These data suggest a nuclear role that may be associated with mRNA processing and/or export. In addition, the PABP2 orthologues have substitutions in residues that are critical for the recognition of poly(A), suggesting a differential mRNA recognition potential (22).

Since RBPs are essential for the control of gene expression we analyzed through mass spectrometry of *Leishmania infantum* PABPs, the RNA binding proteins especially RRM containing, differentially associated to the PABP complexes. Two were selected for better characterization and named RBP23 and DRBD2 according to their homologues in *T. brucei*. In order to better understand the distinct functions of PABPs, mass spectrometry and protein-protein interaction assays were performed for both, RBP23 and DRBD2, to confirm interactions with the PABP homologues. Here, it was defined that RBP23 interacts directly with PABP1 and

share the same partners, while DRBD2 is more related to PABP2 and a large range of proteins. Both complexes are related to distinct mRNA populations which were identified through RNA-seq. This set of experiments can indicate different mechanisms of mRNA selection for translation initiation as for example in the case of RBP23 that specifically binds to ribosomal proteins mRNAs.

MATERIALS AND METHODS

Plasmid constructs, DNA manipulations

The *L. infantum* genomic DNA from the MHOM/MA/67/ITMAP-263 strain was isolated using DNAzol (Life Technologies) following manufacturer's instructions. The full length *RBP23, DRBD2, PABP2* and *PABP3* and three regions of *PABP1* gene (named RRM34, MLLE1 and MLLE2) were amplified using primers flanked by sites for the restriction enzymes BamHI and HindIII. All PCR fragments were first cloned into the pGEM-T Easy vector (Promega) to prior sequencing and subcloning. For immunoprecipitation assays, the genes *RBP23, DRBD2, PABP2* and *PABP3* were subcloned into the BamHI-HindIII sites of the *Leishmania* expression vector pSPBT1YNEOα (26) modified with a 27 nucleotides extension encoding the HA epitope (YPYDVPDYA), added immediately after the coding sequence and prior to the translation stop codon. The production of HA-tagged PABP1 and TP-SP (Phosphorylation PABP1 mutant) has been previously described (27). For pull-down assays, the genes *RPB23* and *DRBD2* were subsequently subcloned into the same sites of the pET-21a vector (Novagen) to be expressed as radiolabeled proteins *in vitro.* GST-tagged *L. infantum* PABP2, PABP3 and three PABP1 truncations (RRM34, MLLE1 and MLLE2) was generated after subcloning into a modified pGEX4T3 expression vector (GE Healthcare) to be expressed as recombinant proteins in *Escherichia coli* (28). The production of GST only, GST-tagged *L. infantum* PABP1 as well as GST-tagged *L. major* PABP1 N- and C-terminal has been also previously described (27). The identity between the *L. major* and *L. infantum* orthologues is 97% therefore the *L. major* protein was considered adequate.

Parasite growth and expression analysis

Leishmania infantum MHOM/MA/67/ITMAP-263 promastigotes were cultured in Schneider's insect medium supplemented with 10% heat-inactivated Fetal Bovine Serum and 2% hemin at pH 7.2, 25ºC. For the expression analysis late exponentially grown *L. infantum* cultures were harvested and resuspended directly into denaturing SDS-PAGE sample buffer (10% SDS, 1 M Tris-HCl pH 6.8, 50% glycerol, bromophenol blue and 5% 2-β-mercaptoethanol), submitted to 15% SDS-PAGE and then blotting with mouse monoclonal antibodies directed against the HA epitope (Anti-HA monoclonal antibody, 100 ng ml−1, Applied Biological Materials). Transfection procedures used for circular plasmids (episomal expression) were carried out by electroporation. Briefly, grown cells were harvested and washed once in HEPES-NaCl buffer (21 mM HEPES pH 7.05, 137 mM sodium chloride, 5 mM potassium chloride, 0.7 mM disodium phosphate, 6 mM glucose) and then transferred to cuvettes containing the plasmid DNA, which were submitted to a pulse of 450 V, 500 µF on the Gene Pulser Xcell^{TM} electroporation system (Bio-Rad). Cells transfected with the *pSPBT1YNEOα* constructs, described below, were selected with G418 (20 µg/ml, Sigma).

Cytoplasmic extract preparation

For RNA sequencing, total cytoplasmic extracts of wild-type *Leishmania major* were produced in triplicate from late exponentially grown cultures of promastigotes cells harvested and washed once in ice cold PBS (Phosphate-buffered saline) and resuspended in IPM1 buffer [100 mM KCl, 5 mM $MgCl_2$, 10 mM HEPES, protease inhibitors (Roche), RNAseOUT and 0.5% IGEPAL® CA-630 (Sigma)] to a concentration of 1-2x10^9 cells/ml. To improve the cells lysis, the resuspended cells were left 10 minutes on ice and then submitted to centrifugation as described above. Additionally, total cytoplasmic extracts from wild-type *L. infantum* and strains expressing HA-tagged RBP23 and DRBD2 as well as HA-tagged PABP1, PABP2 and PABP3 were generated in triplicate after lysing the cells through nitrogen cavitation as described below.

For the mass-spectrometry analysis, total cytoplasmic extracts from recombinant strains of *L. infantum* expressing HA-tagged PABP1, PABP2, PABP3 and a phosphorylation PABP1 mutant (TP-SP) were obtained in duplicate after cells

lysis using through glass beads, acid-washed 425-600 µm *(Sigma)*. Then, total cytoplasmic extracts from wild-type *L. infantum* and recombinant strains expressing HA-tagged RBP23 and DRBD2 were obtained in duplicate after cell lysis through nitrogen cavitation. First, late exponentially grown *L. infantum* promastigotes were harvested and washed once in ice cold PBS, followed by resuspension in HEPES-lysis buffer (20 mM HEPES-KOH pH7.4, 75 mM potassium acetate, 4 mM magnesium acetate, 2 mM DTT, supplemented with and 1x EDTA-free protease inhibitors from Roche) to a concentration of 1-2x10^9 cells/ml. Lysis was carried out using glass beads as described (27) or using nitrogen cavitation as described (29) but with modifications. Briefly, the resuspended cells were transferred into the cavitation chamber of the cell disruption vessel (Parr Instruments) and incubated at 4 ºC under pressure (70 bar) for 40 minutes, followed by rapid decompression and lysis. The lysates were submitted to centrifugation for 10 minutes at 17.000 g, 4 ºC to remove cellular debris and the supernatants, the cytoplasmic extracts, aliquoted and stored at -80 ºC.

Immunoprecipitation

Immunoprecipitation assays (IPs) for the native PABPs used the cytoplasmic extracts from wild type *L. major* and the affinity purified antibodies (anti-PABP1, anti-PABP2 and anti-PABP3) previously described (22) with PBS and RNAseOUT (Sigma) were incubate with approximately 0.1 mg protein A sepharose, previously washed with PBS, overnight at 4 °C. Next day, the sepharose-antibody was washed three times with PBS before incubation with 1x10^9 of cytoplasmic extract for 1 h at 4 °C under agitation. The supernatant was removed by centrifugation for 2 minutes at 600 g, 4 °C and the set were washed three times sequentially with lysis buffer, 1% IGEPAL and lysis buffer. This IP assay using polyclonal antibodies and *L. major* extract was performed exclusively for RNA-seq.

For the IPs with the HA-tagged proteins, all cytoplasmic extracts from wild type or recombinant HA-tagged *L. infantum* strains, were mixed with Pierce™ Anti-HA Magnetic Beads as per manufacturer's protocol. Briefly, 0.2 mg of these anti-HA magnetic beads were washed three times with PBS followed by the incubation with the cytoplasmic extract for 1 h at 4 ºC. After, the depleted supernatant was removed and the beads were washed three times with PBS. The resulting, specifically bound,

immunoprecipitated antigen-antibody complexes were eluted in SDS-PAGE sample buffer. They were then analysed through SDS-PAGE and Western-blotting using antibodies against the HA-tag to confirm the efficiency of the precipitation reaction. The immunoprecipitations were performed in duplicates and triplicates for mass spectrometry and RNA-seq, respectively.

Mass-spectrometry analysis

For mass-spectrometry (MS), IP of HA-tagged PABP1, TP-SP, PABP2 and PABP3 were submitted to Proteomics Platform of the Quebec Genomics Center, as described (30). The results of two independent experiments were analyzed in the Scaffold proteome software, used to validate the proteins identification based in the *L. infantum* genome. Only proteins identified with >1 peptide and a probability of >80.0% were considered. In Table 1 was listed a comparison between proteins co-purified with HA-tagged proteins that were RNA or translation related. Eluted proteins from IP of HA-tagged RBP23 and DBRD2 were submitted to Proteomics facility at the Carlos Chagas Institute - Fiocruz. The samples were loaded into 15% SDS-PAGE gels and allowed to migrate into the resolving gel, when the electrophoresis was interrupted prior to protein fractionation. Gel slices containing the whole IP products were then excised and submitted to an in-gel tryptic digestion and mass spectrometry analysis and validation as previously described (31). Protein identification was based on the *L. infantum* protein sequence database (*L. infantum* JPCM5, version from March 29, 2016 available at TriTrypDB). To confirm the specificity of the IP assays, for each polypeptide, the ratio between the intensity generated from the IPs using the extracts expressing the HA-tagged RBP23 or DRBD2 and the intensity from the control IP using an extract from non-transfected cells was first determined. The base 2 logarithms of the ratio were then calculated for two independent experiments carried out with different cytoplasmic extracts and only the values >3.0 were considered. Identified polypeptides were then ranked and listed in Table 2 and 3 with the highest-ranking values obtained for the 1st experiment listed on top.

RNA extraction and cDNA library construction

RNA ligands were extracted from three independent immunoprecipitations of wild-type *L. infantum*, HA-RBP23, HA-DRBD2 and HA-tagged PABP1-3 with the RNeasy Mini Kit (QIAGEN). The RNA samples were quantified by Qubit™ RNA HS Assay Kit (Thermo Fisher) using to read the concentration the Qubit® 2.0 Fluorometer. At least 0.1-4 µg of the total RNA was used to construct the cDNA library with the TrueSeq Stranded mRNA Library Prep Kit (Illumina). First, the mRNAs were purified by poly(A) and fragmented to synthesize the first strand of cDNA. After that, the second strand of cDNA was synthesized and the 3' ends were adenylated to the ligation of adapters. The DNA fragments that have adapter molecules on both ends were enriched by PCR and then the library was validated quantitatively for qPCR use the KAPA Library Quantification Kits and qualitatively by visualization in agarose gel. Finally, the libraries were normalized and the pool prepared to sequencing using the MiSeq® Reagent Kit v3, 150 cycle (Illumina). Similar RNA extraction was performed for *L. major* cytoplasmic extract IPs with anti-PABP1-3. Then, the total RNA was used to prepare the cDNA library with the SOLiD Whole Transcriptome Analysis Kit and the products were evaluated with an Agilent Bioanalyzer (Agilent). The cDNA library was used to clonal amplification according to the SOLiD Full-Scale Template Bead preparation protocol and sequenced with the SOLiD4 System (Applied Biosystems).

Sequence analysis

Sequence analysis and alignment of the RBP23 and DRBD2 were carried out using the MAFFT version 7 (https://mafft.cbrc.jp/alignment/software/), using selected trypanosomatid sequences recovered from TriTrypDB (https://tritrypdb.org/tritrypdb/) and sequences from other organisms from GenBank. TriTrypDB accessions: *Leishmania infantum* (Li) RBP23 - LinJ.17.0610; *Leishmania major* (Lm) RBP23 - LmjF.17.0550; *Leishmania braziliensis* (Lb) RBP23 - LbrM.17.0540; *Crithidia fasciculata* (Cf) RBP23 - CFAC1_280015000; *Trypanosoma brucei* (Tb) RBP23 - Tb927.10.11270; *Trypanosoma cruzi* (Tc) RBP23 - TcCLB.507711.40; *Bodo saltans* (Bs) RBP23 - CUG90143.1; LiDRBD2 - LinJ.35.2240; LmDRBD2 - LmjF.352200; LbDRBD2 - LbrM.34.2130; CfDRBD2 - CFAC1_300070700; TbDRBD2 - Tb927.9.13990; TcDRBD2 - TcCLB.510755.120; BsDRBD2 - CUG88059.1. Within

the aligned regions, amino acids identical in more than 60% of the sequences are highlighted in dark gray, while amino acids defined as similar, based on the BLOSUM 62 Matrix, on more than 60% of the sequences, are shown in pale gray. The secondary structure was predicted using the Phyre2 (32) automatic fold recognition server and the identification of the arginine methylation sites was done through visual inspection of sequences. The 3' UTR analysis were performed through MEME (33) using 300 nucleotides immediately after the stop codon from the 20 top-most mRNA sequences bound to *L. infantum* RBP23-HA. The already defined 3' UTRs of *T. brucei* orthologues were also retrieved and analyzed. Default parameters were used to detect the motifs, which any number sites per sequence and width between 6-50 nucleotides.

RNA sequencing analysis

RNA-seq reads obtained from SOLiD data were mapped against the *Leishmania major* Friedlin genome assembly version 8.1 available at the TriTrypDB database by the SHRiMP software version 2.2.3 (34) with default parameters. All mappings whose score was higher than 350 were considered for further analyses, where all samples were normalized and the differential expression was assessed by the edgeR package (35), included in the Bioconductor package version 2.6 (36). Genes were considered over represented when a 4-fold increase was observed when compared to the negative control (log_2 ratio >=2), with an FDR<=0.01.

The data from Illumina were analyzed by the following bioinformatics tools: (1) FastQC to evaluate the quality of the sequences (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/); (2) Trimmomatic 0.36 version (37) to remove the adapters and the low-quality sequences; (3) STAR (38), an aligner to map the reads to the *L. infantum* genome, counting reads associated for each gene; (4) DEseq2 Galaxy version (39) to compare statically the samples. The minimum of 4-fold increase over the negative control (Log2 (I Ratio) >2) and FDR of 0.05 were considered. The results of both analyses were plotted on graphs according to gene ontology (GO) functional terms classification of the upregulated mRNAs associated to each protein. The terms/categories used for the grouping were: binding, with a subcategory named DNA binding, which was divided in histone and

non-histone; catalytic activity; structural molecule activity, but subdivided in structural constituent of ribosome or structural constituent of nuclear pore; transcription regulator activity; translation regulator activity and transporter activity.

In vitro pull-down assays

Pull-down assays were performed using Glutathione Sepharose 4B beads (GE Healthcare), affinity purified GST-tagged recombinant proteins and ^{35}S-labeled RBP23 and DRBD2 as described previously (25). GST alone and the different GST-tagged PABP1-3 and PABP1 variants (wild-type and truncations) were expressed from their respective genes cloned in pGEX4T3 in *E. coli*, immobilized on the beads and incubated with ^{35}S-labeled RBP23 and DRBD2. The labeled protein was obtained through the linearization of the respective construct in the pET21a with HindIII, followed by transcription with T7 RNA polymerase in the presence of the cap analogue and translation in the rabbit reticulocyte lysate (Promega or Ambion) supplemented with ^{35}S-methionine (Perkin Elmer). The signal was detected by an autoradiographic film on the 15% polyacrylamide gel.

RESULTS

mRNAs differentially bound to native *Leishmania* PABPs

Immunoprecipitation (IP) assays have previously shown that *Leishmania* PABP1 does not co-precipitate with either PABP2 or PABP3 while, in contrast, reciprocal assays have confirmed that nearly all PABP2 and PABP3 co-precipitate together. These results suggest that PABP1 binds to mRNA targets distinct from the PABP2/PABP3 pair (22). In *T. brucei*, where PABP3 is missing, their PABP1 and PABP2 orthologues have also been shown to bind to different mRNA targets, with PABP1 binding to a subset of small mRNAs, while PABP2 binds to the bulk of the parasite mRNAs (40). Here, to better define the functional distinctions between the *Leishmania* PABPs their mRNAs targets were investigated. To accomplish this we first opted to use previously described rabbit polyclonal sera generated against the three *Leishmania* PABPs and affinity purified using the corresponding GST-tagged recombinant protein (22). These were used to precipitate the corresponding native proteins using protein-A sepharose beads and cytoplasmic extracts derived from *L.*

major promastigotes lysed in the presence of the anionic detergent IGEPAL CA-630. Bound mRNAs were then extracted and submitted to next generation sequencing using the SOLiD platform. Transcripts with the highest number of reads and a $\log_2$ fold enrichment greater than 2 when compared with the negative control (beads incubated with *L. major* cytoplasmic extract without antibodies) were grouped according to GO functional terms (Figure 1A). For these analyses we opted not to include the hypothetical proteins and/or those with unknown functions, although the number of different mRNAs classified as such is shown in the figure.

As expected, the PABP1 profile of bound mRNAs is noticeably different from either PABP2 or PABP3, with the profiles from these last two proteins being very similar (Figure 1A). A total of 216 upregulated mRNAs were found co-precipitated with PABP1, of which only 17 genes encode proteins with unknown functions. Among those transcripts functionally classified, 120 (60%) encode ribosomal proteins followed by 45 (23%) encoding binding proteins of which 17 (9%) are histones mRNAs; 17 (9%) protein-encoding with catalytic activity; 9 (5%) with translation regulator activity; and 8 (4%) with transporter activity. In contrast PABP2/PABP3 comprises a total of 222 and 188 enriched genes which of 26 and 21 with unknown function, respectively. For PABP2, the mRNAs of ribosomal structure encountered about 70 (36%). About 39 (20%) genes encoding proteins with catalytic activity and 58 (30%) with binding function. Notice that about 29 (15%) is histones mRNAs. In addition, 6% of genes encoding translation regulator activity; and 9% with transporter activity were also found to be enriched. The observed data for PABP3 were very similar to those found for PABP2. Transcripts encoding proteins with transcription regulator function were not found in none of the three PABPs.

mRNAs bound to HA-tagged *Leishmania* PABPs

The polyclonal nature of the antibodies used for the set of IPs carried out for the native proteins might be associated with some degree of cross-reactivity and this might be the reason for the overlap seen in mRNAs bound by PABP1 and the PABP2/PABP3 pair. To clarify this possibility, we opted to carry out an independent confirmatory experiment for the analysis of PABP-bound mRNAs. With the availability of a previously described *L. infantum* cell line expressing the HA-tagged PABP1 (27), we now also generated cell lines expressing both PABP2 and PABP3 fused with an

identical C-terminal HA tag. When compared with the ectopic PABP1-HA, represented by two *major* isoforms indicative of phosphorylation events, both HA-tagged PABP2 and PABP3 were visualized as single bands in whole cellular extracts derived from the transfected cell lines (data not shown). Cytoplasmic extracts of these cells lines were then prepared after lysis through cavitation, in the absence of any detergent, and then used in IPs performed with monoclonal anti-HA antibodies immobilized on magnetic beads. The negative control used was a cytoplasmic extract without expressing HA-tagged protein. The co-purified mRNAs were extracted and used for sequencing by the Illumina platform. The new approach then uses different lysis methods as well as distinct antibodies and immunoprecipitation/sequencing strategies. A first analysis of the bound mRNAs with a volcano plot confirms a reduced overlap in mRNA content bound between PABP1xPABP2 or between PABP1xPABP3 (Supplementary Figure 1A). In contrast, no difference between the mRNAs associated to PABP2 or PABP3 can be detected (Supplementary Figure 1B and 1C). This analysis confirms *major* distinctions in mRNA targets bound to PABP1 or PABP2/PABP3, while highlighting once again that the latter two proteins bind together to the same set of mRNAs.

An analysis using GO terms, as carried out for mRNAs immunoprecipitated with the native PABP, was also performed with the new data. A total of 113 upregulated mRNAs were found co-precipitated with PABP1 fulfilling the criteria of $\log_2$ fold enrichment greater than 2, of which 27 transcripts encode proteins with unknown functions (Figure 1B). From the remaining transcripts, 50 (58%) are grouped within the term structural constituent of ribosome, 19 (22%) encode binding proteins, 13 (15%) enzymes with catalytic activity, 2 (2%) proteins with transporter activity; 1 (1%) one each encoding proteins classified as transcription and translation regulator activities. PABP2 and PABP3 co-precipitated with totals of 135 and 118 enriched transcripts, respectively, of which 48 and 40 with unknown functions. Among the mRNAs associated with PABP2 and PABP3, only 3% and 4% were classified within the “structural constituent of ribosome” term. For PABP2, 35 (40%) transcripts encode proteins with catalytic activity and 38 (43%) with binding function. It is notable that the mRNAs co-precipitated with HA-tagged PABP1, PABP2 and PABP23 are not so enriched with histone mRNAs as compared with *L. major*. In addition, 8 (9%) genes encoding proteins with transporter activity, 2 (2%) with

translation regulator activity and 1 (1%) transcription regulator were also found to be enriched. The observed data for PABP3 were very similar to those found for PABP2. Although the results resemble the first experiment with native proteins with respect to PABP1 to have a different profile of PABP2 and PABP3 which have a very similar profile to each other. The experiment with HA-tagged proteins proved to be much more specific, as for PABP1 which clearly has a preference for mRNAs encoding ribosomal proteins.

RBPs differentially bound to PABP homologues *in vivo*

Aside from their strong specificity for poly(A) sequences (22, 41), so far the best studied eukaryotic PABPs have no known sequence specific affinity which could justify their selective binding to different sets of mRNAs. Binding to poly(A) has been shown to be mediated by specific residues within the PABP's RRMs 1 and 2 which are generally conserved in *Leishmania* PABP1 and PABP3, although substitutions have been identified for PABP2 which might lead to changes in sequence binding specificity (22). For *Leishmania* PABP1 at least, and most likely for the PABP2/PABP3 pair also, their recognition of specific mRNAs targets might require the assistance of partner RBPs with distinct RNA binding specificities. Here, putative RNA binding proteins that might be involved in specific mRNA recognition were initially identified in a pilot mass spectrometry analysis of the proteins co-precipitated with the three different, HA-tagged, *L. infantum* PABPs. For these experiments no quantitative data was derived from the mass spectrometry analysis; however, a minimum of two independent IPs was considered for validation. PABP1, or a previously described PABP1 phosphorylation mutant (TP-SP) (27), was thus compared with PABP2 and PABP3. As described elsewhere, the data for the wild-type PABP1 was subsequently confirmed using the milder lysis method based on cavitation followed by a quantitative mass spectrometry analysis (27). These experiments, summarized in the Table 1, mainly confirmed the strong association between PABP2 and PABP3, and with similar binding partners, as well as the specific association of the EIF4E4 initiation factor with PABP1. They also led to the identification of RNA binding proteins specifically associated with PABP1 (RBP23; DRBD3), PABP2/PABP3 (DRBD2; NRBD) or all three proteins (ZC3H41).

Selection of RBP23 and DRBD2 and expression in *Leishmania infantum*

Considering a possible role of RBPs in assisting the binding of the their PABP partners to specific mRNAs, we then chose the two of them more strongly associated with either PABP1 or the PABP2/PABP3, respectively RBP23 and DRBD2, for further investigation. These two RBPs are both characterized by the presence of RRM domains, with one RRM previously described for RBP23 and two for DRBD2 (42). RBP23 is a 32.3 kDa protein (295 amino acids long) having its described RRM localized at its C-terminal end (Figure 2A and Supplementary Figure 2). Here, a second atypical RRM-like domain was identified by the Phyre2 program based on secondary structure predictions and localized to its N-terminal end. DRBD2 is also a small protein with a molecular weight of 30.2 kDa (274 amino acids long) with its two RRMs localized to the N and C-terminal ends of the protein Figure 2B and Supplementary Figure 3). The two proteins therefore are very similarly organized.

Both RBP23 and DRBD2 genes were cloned and used for the ectopic expression of the C-terminally HA-tagged proteins in transfected *L. infantum* cell lines. As for the PABPs, whole cell extracts of these proteins were first tested with anti-HA antibody by western-blot assays. The expression of these proteins was confirmed in exponentially grown promastigotes, with RBP23 and DRBD2 migrating as single bands with predicted sizes in agreement with the proteins' molecular weights (Figure 2C). No isoforms suggestive of post-translational modifications were seen. Prior to subsequent analyses investigating protein partners and mRNA targets, cytoplasmic extracts were then generated from the cell lines expressing both HA-tagged proteins and preliminary IPs were performed. The amount of RBP23 detected in the cytoplasmic extract and/or IPs is generally much lower than that observed for DRBD2 and contrasts with equivalent levels of expression for both ectopic proteins from whole cell extracts. RBP23 then is likely to be more susceptible to degradation than DRBD2.

RBP23 binding partners

In *T. brucei*, RBP23 has been shown to stimulate translation when tethered to the 3' end of a reporter mRNA, a profile also seen for both PABP homologues (7, 43). As expected, it is localized in cytoplasm with a reticulated distribution (44). RNAi mediated depletion of RBP23 affects growth of the *T. brucei* bloodstream form (45)

and the cells show a gain-of-fitness phenotype during differentiation, but depletion does not seem to affect procyclic cells (46). In *L. infantum* studies with isobaric tagging methodology showed RBP23 to be significantly downregulated during differentiation and in the mature amastigote form (47). Here, an alignment comparing the sequences of RBP23 orthologues from five trypanosomatids plus the free-living relative *Bodo saltans* was carried out in order to find conserved or divergent features within the protein's sequence (Supplementary Figure 2). The alignment reveals a limited conservation in sequence which is mainly seen between more closely related species, such as within the *Leishmania* genus, and much more restricted between more distantly related organisms, such as *B. saltans*. When the *L. infantum* and *T. brucei* RBP23 orthologues are compared with 100% of coverage, only 30% identity and 44% of similarity in sequence is seen, for instance, with the conservation mainly restricted to the RBP23 N- and C-terminal regions, where the atypical and typical RRMs have been mapped. Apart from the RRMs, some conservation is also observed in the segment connecting them with some aromatic residues. Noteworthy also is the presence of a glutamine rich sequence followed by a tract of multiple prolines found only in the *Leishmania* RBP23 orthologues as well as several RGG motifs present only in the *Leishmania* and *Crithidia* sequences. RGG motifs have been implicated in arginine methylations in trypanosomatids, with a role in regulation of gene expression and cell cycle (48).

To confirm the association of RBP23 with PABP1 and find additional protein partners, proteins co-precipitated with the HA-tagged RBP23 were next identified through mass-spectrometry. Results derived from the mass spectrometry analysis of two sets of replicates comparing the LFQ intensity of RBP23-HA with the respective negative control are summarized in Table 2. These results confirm the specific interaction between RBP23 with PABP1 as well as the translation initiation complex formed by EIF4E4 and EIF4G3, the top three associated proteins. Several other proteins which have been shown to specifically co-precipitate with HA-tagged PABP1 using similar procedures (27) are also found associated with RBP23-HA. These include the zinc-finger ZC3H41, structurally similar to a DNA helicase according to Phyre2 predictions and has also been found associated with polysomes (49)**.** Also associated with both PABP1 and RBP23 is the hypothetical protein LinJ.18.0300, having a Nuclear Transport Factor 2 (NFT2)-like domain in its N-terminus. Yet

another protein co-precipitated with both PABP1 and RBP23 is LinJ.05.0450, structurally similar to Skp1 (S-phase kinase associated protein 1), a ubiquitin ligase complex component. Two other hypothetical proteins, with unknown function, were found co-precipitated with RBP23 but were not with PABP1: LinJ.04.0140 and LinJ.19.0020. RBP23, but not PABP1 was also found associated with a potential Pentatricopeptide Repeat (PPR) protein, rich in alpha helixes and which in several organisms are involved with RNA processing (50)**.** At last RBP23 co-precipitated with two subunits of the CCR4-NOT transcription complex, NOT1, a scaffolding protein, and CAF1, which displays deadenylase activity. This complex is an important regulator during all gene expression pathway, interacting with transcription and translation activator and repressor factors and promoting mRNA decay (51)**.**

The orthologues in *T. brucei* of the six partner proteins are all cytoplasmic with Tb927.7.7460 (LinJ.05.0450) and Tb927.9.7080 (LinJ.04.0140) exclusively in the cytoplasm, ZC3H41 also in nucleoplasm, Tb927.10.13800 (LinJ.18.0300) also in axoneme, Tb927.10.14700 (LinJ.19.0020) in nucleoplasm and nucleolus, and the Pentatricopeptide Repeat Protein (PPR) also in flagellar cytoplasm and nuclear lumen (44). RNAi-induced knockdown promotes a loss-of-fitness in three orthologues of RBP23 partners: Tb927.10.13800 (LinJ.18.00300) only in procyclic; Tb927.7.7460 (LinJ.05.0450) in bloodstream and during differentiation; and Tb927.9.7080 (LinJ.04.0140) and ZC3H41 in both forms and also during differentiation. The RNAi-induced knockdown in Tb927.14700 (LinJ.19.0020) and in PPR does affect either procyclic, bloodstream or the differentiation between both forms (46)**.**

DRBD2 binding partners

In contrast to RBP23, DRBD2 has been shown to repress translation when tethered to a reporter mRNA (7, 43)**.** In procyclic, no increase in ZC3H11, a stress-response RNA-binding protein, was observed during DRBD2 RNAi, which means that possibly the lack of DRBD2 does not disturb the cell so that ZC3H11 is not required (52). This protein is located in the cytoplasm distributed irregularly or in the form of points with more intense signal near the nucleus (44)**.** In *L. infantum*, different DRBD2 isoforms were expressed in both amastigote and promastigote stages, but were found to interact with *L. infantum* Alba1 and Alba3 specifically only in amastigotes form through mass spectrometry analysis (30, 53). The alignment of

several DRBD2 homologues (Supplementary Figure 3) demonstrated that the sequence is highly conserved when compared with RBP23, especially in the RRMs. For example, comparing *L. infantum* and *T. brucei* DRBD2 homologues with 91% of coverage was observed 62% of identity and 70% of similarity. However, *T. brucei* reveals a polyglutamine (poly Q) between the two RRMs, not conserved in *L. infantum.* Interestingly, this specie presents an element with aromatic residues between the RRMs, as observed to RBP23.

Table 3 compares the results derived from the mass spectrometry analysis of two sets of replicates comparing the LFQ intensity of DRBD2-HA with the respective negative control. The data from Table 3 confirms that DRBD2 preferentially associates with PABP2 and 3 as previously demonstrated. Curiously, DRBD2 also interacts with the translation initiation factor EIF4G3 involved in the complex associated to PABP1. However, most of the proteins that interact with DRBD2 are RNA related. In our analyses a total of four Zinc Finger proteins, two RBPs, two PUFs and two TRRMs were found. Highlight for Pumilio protein 6 which is a translation repressor in *Saccharomyces cerevisiae* (54) and is involved in retroposon-mediated mRNA decay in *Leishmania* (55)**.** We can also observe nine mitochondrial RNA related proteins, four of which are associated with guide RNA, known as group of mitochondrial RNAs with role in mRNA editing in *T. brucei* (56)**,** and the polyadenylation proteins, KPAF1 and PAMC1. In addition, DBRD2 was also associated with three enzymes, one involved in glycolysis and the other two with RNA helicase activity. The set of proteins associated with DBRD2 demonstrates that it is somehow involved in regulation but not directly associate to translation, as seen for RBP23.

Mapping the interactions regions between RBP23 and DRBD2 to PABP1 *in vitro*

The PABP RMMs 1 and 2 are responsible for the specific binding to the poly(A) tail of the mRNAs and also mediates the PABP interaction with eIF4G factor (57–59). Less is known regarding the functional roles of RRMs 3 and 4, however they have been shown to be able to bind to non-polymeric AU sequences, for unknown reasons, and also were seen to mediate the interaction with other PABP partners such as the elongation factor 1 alpha (60–62). The role for the MLLE C-terminal domain in mediating protein-protein interactions has been more clearly defined and it

has been shown to mediate the interaction between PABP and several different proteins such as the translation termination factor eRF3 and the mammalian regulatory proteins Paip1 and Paip2 (19, 63, 64). Indeed, the *Leishmania* PABP1 MLLE has also been shown to mediate its interaction with EIF4E4 (27, 65)**.** Here, to confirm if direct interactions occur between the RBPs selected here, RBP23 and DRBD2, with the *Leishmania* PABP homologues, *in vitro* pull-down assays were carried out. The three PABPs were first expressed in *E. coli* with an N-terminal Glutathione S-transferase (GST) tag. After immobilization in Glutathione Sepharose these were then incubated with ^{35}S-labeled RBP23 and DRBD2, produced by *in vitro* transcription/translation. The two ^{35}S-labeled proteins bound to all three GST-tagged PABPs, although a stronger signal with PABP1 was observed (Figure 3B). No binding to the negative GST control was observed.

Both ^{35}S-labeled RBP23 and DRBD2 were next tested with GST-tagged deletion variants of PABP1 described in Figure 3A. In this experiment, the ^{35}S-RBP23 bound to the C-terminus construct. There were also a weaker signal of interaction with RRM34 but none with MLLE (data not shown). The ^{35}S-DRBD2 bound to RRM34 (Figure 3C) but in data not shown there was also a weaker signal of binding with N and C-terminus. This could indicate that the interaction occurs preferentially through the RRM3, due to a small region of overlap between N- and C-terminal.

mRNAs populations associated with RBP23 and DRBD2 are different

To understand the function of the complexes PABP1/RBP23 and PABP2/DRBD2 in *Leishmania,* similar assays were performed for HA-tagged RBP23 and DRBD2, with the intention to compare analyses from Illumina data. For this experiment, mRNAs bound to either RBP23-HA or DRBD2-HA after the IPs were extracted and used to construct cDNA libraries and to perform next generation sequencing as was done for the PABPs. The data showed the mRNAs bound to both RBP23-HA and DRBD2-HA. The upregulated mRNAs were functionally grouped according their molecular function description, as previously described for the HA-tagged PABPs. A total of 132 mRNAs enriched were found co-precipitated with RBP23 of which 8 genes encode proteins with unknown functions. The most enriched mRNA was the RBP23 transcript itself and the *major*ity was 112 (90%) genes encoding ribosomal proteins, much more specific than PABP1 with about 60%

of ribosomal proteins mRNAs (Figure 1A). In addition, 7 (6%) genes encoding binding proteins (including RBP23 gene); 3 (2%) protein-coding gene with transcription regulator activity; 1 (1%) with translation regulator activity, the initiation factor 5a; and 1 (1%) with transmembrane transporter activity, a pteridine transporter, were also found to be enriched (Figure 4A).

A different profile was observed for DBRD2 with a total of 301 mRNAs found co-precipitated of which 88 encode proteins with unknown functions. The two largest groups of mRNAs associated to DRBD2 comprises 94 (44%) genes encoding binding proteins and 68 (32%) with catalytic activity. In contrast to RBP23, only 18 (8%) genes encoding proteins with structural molecule activity (16 constituent of ribosome and 2 of nuclear pore) were found enriched with DRBD2. In addition, 18 (8%) protein-encoding gene with transporter activity, 9 (4%) with translation regulator and 6 (3%) with transcription regulator activity were also found to be enriched (Figure 4B).

In silico motifs on RP mRNAs 3'UTR

Next, we hypothesize about the sequence elements of the mRNAs, especially for RP mRNAs, that allow them to be selected by RBP23, since it is known that functionally related transcripts usually have common motifs for the binding of specific RBPs (66). In many mammals, ribosomal protein mRNAs share similar 5' UTRs with several pyrimidines known as TOP elements responsible for the regulation of their translation (67, 68). In trypanosomatids, a mini-exon containing between 39 and 41 nucleotides with a modified cap structure, referred as spliced leader is added in 5' end during trans-splicing (69)**.** Besides, their mRNAs 5′ end is not enriched in pyrimidines, so regulation of their translation must differ and be preferentially located in the 3' UTR of their transcripts (70). Considering the finding that RBP23 binds mostly to mRNAs encoding ribosomal proteins, we investigated conserved motifs within the 3' untranslated regions of these mRNAs from both *L. infantum* and *T. brucei*. The 5' UTRs of these mRNAs are very small and have no evidence of a regulatory sequence as in mammals. Then it was compared the 3' UTR of 19 sequences of ribosomal protein mRNAs more enriched than control, as well as the RBP23 mRNA sequence which was also highly enriched in the analyzes. A motif composed by T-rich sequences of 15 bp in *L. infantum* (TTTHYTTTTKTTTTB) (Figure 5A) and also in *T. brucei* (TBTBTTYTTYYTTTS) (Figure 5B) was found one

or even four times in most 3' UTRs of ribosomal proteins mRNAs analyzed but was not detected in the 3' UTR of RBP23 mRNA. In *T. brucei* the ribosomal protein mRNAs motif was found mostly at the end of already known 3' UTR (data not shown). The UTRs of *L. infantum* are not yet defined, but a large part of the motifs was found before the 150 pairs after stop codon (Figure 5C).

DISCUSSION

Our data confirm the association of RBP23 with one EIF4F complex. This association with known translational activators as PABP1, EIF4E4 e EIF4G3 may suggest a common function involving ribosome recruitment during translation initiation (27). In *T. brucei,* RBP23 interacts exclusively with PABP1, which also interacts with Tb27.7.7460 and ZC3H41 (40) both were found in our mass spectrometry analysis of RBP23. PABP2 interacts with DRBD2, and also with several other proteins, as EIF4G3, ZC3H39, ZC3H40 and PUF6 (40), also found in our DRBD2 analysis. Furthermore, both RBP23 and DRBD2 together with PABPs and Alba 2 protein were co-purified with TSR1 which is involved in splicing regulation, mRNA stability, and rRNA processing (71). The data highlights the importance of these proteins in the mRNA regulation process.

Here, we also confirmed that the both RBPs are capable of binding to all three PABPs, which possibly have a similar interacting protein domain. However, this does not necessarily occur *in vivo*, since the pull-down assays may be missing elements that may disturb specific protein-protein interactions. Human RRM1 and RRM2 are required for the specific recognition of poly(A) tracts, whereas RRM3 and RRM4 can associate with any RNA. Humans RRM3 and 4 bind poly(A) with reduced affinity, but also bind AU-rich RNA and mediate protein-protein interactions (15)**.** Our assays demonstrated that both RRM3 and 4 are involved in the interaction of DRBD2 and RBP23, respectively, with PABP1. None of the two RBPs bound to the MLLE domain only, shown to bind EIF4E4, a known PABP1 partner (25).

Analysis of the main set of transcripts selected by RBP23 and DRBD2 revealed that each might interact with a specific population of mRNAs. The distinctions for different target mRNAs may result from the probable differential function between their PABP partners. PABP1 has been found to be strongly

associated with the essential EIF4E4/EIF4G3 complex, with a predicted relevant role during protein synthesis of most mRNAs (22, 72). RBP23 may help to recruit PABP1 and the EIF4E4/EIF4G3 complex to selected mRNAs to promote their translation (27) which explains the similarity of ribosomal transcripts population found associated with PABP1 and RBP23. In yeast, RNA immunoprecipitation sequencing analysis demonstrated that many ribosomal protein (RP) mRNAs are enriched specifically with the closed loop translation initiation components, eIF4E and both isoforms of eIF4G (73)**.**

In contrast to PABP1, PABP2 (*T. brucei* and *L. major*) and PABP3 (*L. major*), migrate to the nucleus upon inhibition of transcription by Actinomycin D (22, 23) which may indicate a possible nuclear function. As in *T. brucei* DRBD3, a stabilizing protein, which is localized mainly in the cytoplasm but under oxidative stress remains bound to mRNA and concentrates in the nucleus, a typical behavior of a protein that transports target mRNAs (74). Furthermore, PABP2 contains several substitutions in residues implicated in poly(A) recognition, suggesting changes in specificity for the recognition of target mRNAs (22, 75). Thus, the transcripts associated are expected to be different from those of PABP1 and more similar to those found associated with RBP partners, being DRBD2 a potential one. In *T. brucei*, DRBD2 was found distributed accumulated in perinuclear region (44) then the next step will be to evaluate these proteins in presence of inhibitors. Due the localization of DRBD2, we believe that maybe it will migrate to nucleus as well as PABP2, correlating the function of both proteins.

Besides ribosomal proteins, some translation factors are also encoded by mRNAs with TOP elements immediately before the cap structure. LARP1 associates to TOP motif and the cap structure and act as a repressor of TOP mRNA translation, mTORC1 phosphorylates LARP1 and releases it from the TOP motif. mTORC1 also controls the phosphorylation of 4E-BP1, allowing the eIF4G to bind eIF4E and recruit the 43S complex. The LARP1 also binding to PABP and form a translation-inactive mRNA loop (76, 77). In contrast to RP mRNAs which presents a downstream regulatory mechanism, histones mRNAs are upstream regulated by phosphorylation of the stem-loop-binding protein which binds the stem loop structure in 3’ UTR (78, 79).In eukaryotes, different RNA binding protein binds to cis-acting elements in 3’ UTRs and regulates mRNA translation. This regulation occurs by targeting the mRNA

cap structure, eIF4E-eIF4G complex, ribosomes, and the poly(A) tail. For example, many RBPs bound to 3'UTR of mRNAs can associate with deadenylases complex as CCR4-NOT to enhance deadenylation (80). In *T. brucei*, it was found sequences in 3' UTR that regulate the mRNA degradation as AU-rich elements (5). In bloodstream forms, ZC3H11 is bound to these mRNA elements, and recruits PABP via MKT1 and PBP1, which is a possible mechanism for mRNA stabilization (81).

Surprisingly, RBP23 co-precipitated with its own mRNA. Similar to what happens with *S. cerevisiae* 4E-BP protein as known as the Caf20p (a translation repressor) which also interacts with its own transcript, may representing an auto regulatory mechanism (73). As observed *in vitro* studies for PABP binding to the adenylated tract present in the 5 'UTR of its mRNA, which is necessary and sufficient to represses its own synthesis, supporting a model of autoregulation (82)**.** In our model (Figure 6) we proposed that RBP23 interacts with the 3' UTR motif of the ribosomal protein mRNA and its own mRNA to subsequently interact with PABP1. Thus, RBP23 'tag' the ribosomal protein mRNAs that should be translated by the complex PABP1 / EIF4E4 / EIF4G3 / EIF4AI. Some of the hypothetical proteins found co-precipitated with RBP23 may aid in their function or in the translation process itself.

FUNDING

This work was supported by Science and Technology Support Foundation of Pernambuco (FACEPE) [APQ-0239-2.02/12]; National Council for Scientific and Technological Development (CNPq) [475471/2008-3, 480899/2013-4]; and Coordination for the Improvement of Higher Education Personnel (CAPES) [23038.007656/2011-92].

ACKNOWLEDGMENTS

We thank all the members of Dr. de Melo Neto's laboratories for helpful discussions and technical assistance. The authors thank the Program for Technical Development of Health Inputs-PDTIS-FIOCRUZ for the funding and use of the

automatic sequencing facility at the Instituto Aggeu Magalhães and the proteomics facility at the Instituto Carlos Chagas.

REFERENCES

1. Haile,S. and Papadopoulou,B. (2007) Developmental regulation of gene expression in trypanosomatid parasitic protozoa. *Curr. Opin. Microbiol.*, **10**, 569–577.
2. Michaeli,S. (2011) Trans-splicing in trypanosomes: machinery and its impact on the parasite transcriptome. *Future Microbiol.*, **6**, 459–474.
3. Requena,J.M. (2011) Lights and shadows on gene organization and regulation of gene expression in *Leishmania*. *Front Biosci*, **16**, 2069–2085.
4. Clayton,C. and Shapira,M. (2007) Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. *Mol Biochem Parasitol*, **156**, 93–101.
5. Kramer,S. and Carrington,M. (2011) Trans-acting proteins regulating mRNA maturation, stability and translation in trypanosomatids. *Trends Parasitol.*, **27**, 23–30.
6. Müller-McNicoll,M. and Neugebauer,K.M. (2013) How cells get the message: dynamic assembly and function of mRNA-protein complexes. TL - 14. *Nat. Rev. Genet.*, **14 VN-r**, 275–287.
7. Erben,E.D., Fadda,A., Lueong,S., Hoheisel,J.D. and Clayton,C. (2014) A Genome-Wide Tethering Screen Reveals Novel Potential Post-Transcriptional Regulators in *Trypanosoma brucei*. *PLoS Pathog.*, **10**.
8. Hogan,D.J., Riordan,D.P., Gerber,A.P., Herschlag,D. and Brown,P.O. (2008) Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system. *PLoS Biol.*, **6**, 2297–2313.
9. Cheng,M.H. and Jansen,R.-P. (2017) A jack of all trades: the RNA-binding protein vigilin. *Wiley Interdiscip. Rev. RNA*, **8**, e1448.
10. Oliveira,C., Faoro,H., Alves,L.R. and Goldenberg,S. (2017) RNA-binding proteins and their role in the regulation of gene expression in *Trypanosoma cruzi* and saccharomyces cerevisiae. *Genet. Mol. Biol.*, **40**, 22–30.
11. Clayton,C. (2013) The Regulation of Trypanosome Gene Expression by RNA-

Binding Proteins. *PLoS Pathog.*, **9**, 9–12.

12. Kolev,N.G., Ullu,E. and Tschudi,C. (2014) The emerging role of RNA-binding proteins in the life cycle of *Trypanosoma brucei*. *Cell. Microbiol.*, **16**, 482–489.
13. Najafabadi,H.S., Lu,Z., MacPherson,C., Mehta,V., Adoue,V., Pastinen,T. and Salavati,R. (2013) Global identification of conserved post-transcriptional regulatory programs in trypanosomatids. *Nucleic Acids Res.*, **41**, 8591–8600.
14. Eliseeva,I.A., Lyabin,D.N. and Ovchinnikov,L.P. (2013) Poly(A)-binding proteins: structure, domain organization, and activity regulation. *Biochem. Biokhimiia*, **78**, 1377–91.
15. Goss,D.J. and Kleiman,F.E. (2013) Poly(A) binding proteins: Are they all created equal? *Wiley Interdiscip. Rev. RNA*, **4**, 167–179.
16. Kahvejian,A., Roy,G. and Sonenberg,N. (2001) The mRNA closed-loop model: The function of PABP and PABP-interacting proteins in mRNA translation. *Cold Spring Harb. Symp. Quant. Biol.*, **66**, 293–300.
17. Gallie,D.R. (2014) The role of the poly (A) binding protein in the assembly of the cap-binding complex during translation initiation in plants. *Translation*, **0731**, 1–14.
18. Huntzinger,E., Kuzuoğlu-Öztürk,D., Braun,J.E., Eulalio,A., Wohlbold,L. and Izaurralde,E. (2013) The interactions of GW182 proteins with PABP and deadenylases are required for both translational repression and degradation of miRNA targets. *Nucleic Acids Res.*, **41**, 978–994.
19. Melo,E.O., Dhalia,R., De Sa,C.M., Standart,N. and De Melo Neto,O.P. (2003) Identification of a C-terminal poly(A)-binding protein (PABP)-PABP interaction domain: Role in cooperative binding to poly(A) and efficient cap distal translational repression. *J. Biol. Chem.*, **278**, 46357–46368.
20. Lin,J., Fabian,M., Sonenberg,N. and Meller,A. (2012) Nanopore detachment kinetics of poly(A) binding proteins from RNA molecules reveals the critical role of C-terminus interactions. *Biophys. J.*, **102**, 1427–1434.
21. Kühn,U. and Wahle,E. (2004) Structure and function of poly(A) binding proteins. *Biochim. Biophys. Acta - Gene Struct. Expr.*, **1678**, 67–84.
22. da Costa Lima,T.D., Moura,D.M.N., Reis,C.R.S., Vasconcelos,J.R.C., Ellis,L., Carrington,M., Figueiredo,R.C.B.Q. and de Melo Neto,O.P. (2010) Functional characterization of three *Leishmania* poly(A) binding protein homologues with

distinct binding properties to RNA and protein partners. *Eukaryot. Cell*, **9**, 1484–1494.

23. Kramer,S., Bannerman-Chukualim,B., Ellis,L., Boulden,E.A., Kelly,S., Field,M.C. and Carrington,M. (2013) Differential Localization of the Two *T. brucei* Poly(A) Binding Proteins to the Nucleus and RNP Granules Suggests Binding to Distinct mRNA Pools. *PLoS One*, **8**.
24. Moura,D.M., Reis,C.R., Xavier,C.C., da Costa Lima,T.D., Lima,R.P., Carrington,M. and de Melo Neto,O.P. (2015) Two related trypanosomatid eIF4G homologues have functional differences compatible with distinct roles during translation initiation. *RNA Biol.*, **12**, 305–19.
25. De Melo Neto,O.P., Da Costa Lima,T.D.C., Xavier,C.C., Nascimento,L.M., Romão,T.P., Assis,L. a., Pereira,M.M.C., Reis,C.R.S. and Papadopoulou,B. (2015) The unique *Leishmania* EIF4E4 N-terminus is a target for multiple phosphorylation events and participates in critical interactions required for translation initiation. *RNA Biol.*, **12**, 1209–1221.
26. Padmanabhan,P.K., Dumas,C., Samant,M., Rochette,A., Simard,M.J. and Papadopoulou,B. (2012) Novel features of a PIWI-like protein homolog in the parasitic protozoan *Leishmania*. *PLoS One*, **7**, e52612.
27. de Melo Neto,O.P., da Costa Lima,T.D.C., Merlo,K.C., Romão,T.P., Rocha,P.O., Assis,L.A., Nascimento,L.M., Xavier,C.C., Rezende,A.M., Reis,C.R.S., *et al.* (2018) Phosphorylation and interactions associated with the control of the *Leishmania* Poly-A Binding Protein 1 (PABP1) function during translation initiation. *RNA Biol.*, **15**, 739–755.
28. Dhalia,R., Reis,C.R.S., Freire,E.R., Rocha,P.O., Katz,R., Muniz,J.R.C., Standart,N. and De Melo Neto,O.P. (2005) Translation initiation in *Leishmania major*: Characterisation of multiple eIF4F subunit homologues. *Mol. Biochem. Parasitol.*, **140**, 23–41.
29. Mureev,S., Kovtun,O., Nguyen,U.T.T. and Alexandrov,K. (2009) Species-independent translational leaders facilitate cell-free expression. *Nat. Biotechnol.*, **27**, 747–752.
30. Dupé,A., Dumas,C. and Papadopoulou,B. (2015) Differential subcellular localization of *Leishmania* alba-domain proteins throughout the parasite development. *PLoS One*, **10**, 1–29.

31. Rezende,A.M., Assis,L.A., Nunes,E.C., da Costa Lima,T.D., Marchini,F.K., Freire,E.R., Reis,C.R. and de Melo Neto,O.P. (2014) The translation initiation complex eIF3 in trypanosomatids and other pathogenic excavates--identification of conserved and divergent features based on orthologue analysis. *BMC.Genomics*, **15**, 1175.
32. Kelley,L., Mezulis,S., Yates,C., Wass,M. and Sternberg,M. (2015) The Phyre2 web portal for protein modelling, prediction, and analysis. *Nat. Protoc.*, **10**, 845–858.
33. Bailey, Timothy L.; Elkan,C. (1994) Fitting a mixure model by expectation maximization to discover motifs in biopolymers. In *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology*. AAAI press, Menlo Park, California, pp. 28–36.
34. David,M., Dzamba,M., Lister,D., Ilie,L. and Brudno,M. (2011) SHRiMP2: Sensitive yet practical short read mapping. *Bioinformatics*, **27**, 1011–1012.
35. Robinson,M.D., McCarthy,D.J. and Smyth,G.K. (2009) edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139–140.
36. Gentleman RC,C. V, Bates,D.M., Bolstad,B., Dettling,M., Dudoit,S., Ellis,B., Gautier,L., Ge,Y., Gentry,J., Hornik,K., *et al.* (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol*, **5**.
37. Bolger,A.M., Lohse,M. and Usadel,B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.
38. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
39. Afgan,E., Baker,D., Batut,B., Van Den Beek,M., Bouvier,D., Ech,M., Chilton,J., Clements,D., Coraor,N., Grüning,B.A., *et al.* (2018) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.*, **46**, W537–W544.
40. Zoltner,M., Krienitz,N., Field,M.C. and Kramer,S. (2018) Comparative proteomics of the two *T. brucei* PABPs suggests that PABP2 controls bulk mRNA. *PLoS Negl. Trop. Dis.*, **12**, 1–18.
41. Bates,E.J., Knuepfer,E. and Smith,D.F. (2000) Poly (A) -binding protein I of

Leishmania : functional analysis and localisation in trypanosomatid parasites. *Nucleic Acids Res.*, **28**, 1211–1220.

42. Gaudenzi,J. De, Frasch,A.C. and Clayton,C. (2005) RNA-Binding Domain Proteins in Kinetoplastids : a Comparative Analysis RNA-Binding Domain Proteins in Kinetoplastids : a Comparative Analysis †. *Eukaryot. Cell*, **4**, 2106–2114.
43. Lueong,S., Merce,C., Fischer,B., Hoheisel,J.D. and Erben,E.D. (2016) Gene expression regulatory networks in *Trypanosoma brucei*: Insights into the role of the mRNA-binding proteome. *Mol. Microbiol.*, **100**, 457–471.
44. Dean,S., Sunter,J.D. and Wheeler,R.J. (2017) TrypTag.org: A Trypanosome Genome-wide Protein Localisation Resource. *Trends Parasitol.*, **33**, 80–82.
45. Wurst,M., Robles,A., Po,J., Luu,V.D., Brems,S., Marentije,M., Stoitsova,S., Quijada,L., Hoheisel,J., Stewart,M., *et al.* (2009) An RNAi screen of the RRM-domain proteins of *Trypanosoma brucei*. *Mol. Biochem. Parasitol.*, **163**, 61–65.
46. Alsford,S., Turner,D.J., Obado,S.O., Sanchez-flores,A., Glover,L., Berriman,M., Hertz-fowler,C. and Horn,D. (2011) High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. *Genome Res.*, **21**, 915–924.
47. Rosenzweig,D., Smith,D., Opperdoes,F., Stern,S., Olafson,R.W. and Zilberstein,D. (2008) Retooling *Leishmania* metabolism: from sand fly gut to human macrophage. *FASEB J.*, **22**, 590–602.
48. Fisk,J.C. and Read,L.K. (2011) Protein arginine methylation in parasitic protozoa. *Eukaryot. Cell*, **10**, 1013–1022.
49. Klein,C., Terrao,M., Gil,D.I. and Clayton,C. (2015) Polysomes of *Trypanosoma brucei*: Association with initiation factors and RNA-binding proteins. *PLoS One*, **10**, 1–12.
50. Manna,S. (2015) An overview of pentatricopeptide repeat proteins and their applications. *Biochimie*, **113**, 93–99.
51. Collart,M.A. (2016) The Ccr4-Not complex is a key regulator of eukaryotic gene expression. *Wiley Interdiscip. Rev. RNA*, **7**, 438–454.
52. Minia,I. and Clayton,C. (2016) Regulating a Post-Transcriptional Regulator: Protein Phosphorylation, Degradation and Translational Blockage in Control of the Trypanosome Stress-Response RNA-Binding Protein ZC3H11. *PLoS*

Pathog., **12**, 1–31.

53. Brotherton,M.C., Racine,G., Foucher,A.L., Drummelsmith,J., Papadopoulou,B. and Ouellette,M. (2010) Analysis of stage-specific expression of basic proteins in *Leishmania infantum*. *J. Proteome Res.*, **9**, 3842–3853.
54. Shahbabian,K., Jeronimo,C., Forget,A., Robert,F. and Chartrand,P. (2014) Co-transcriptional recruitment of Puf6 by She2 couples translational repression to mRNA localization. *Nucleic Acids Res.*, **42**, 8692–8704.
55. Azizi,H., Dumas,C. and Papadopoulou,B. (2017) The Pumilio-domain protein PUF6 contributes to SIDER2 retroposon-mediated mRNA decay in *Leishmania*. *Rna*, **23**, 1874–1885.
56. Reifur,L., Yu,L.E., Cruz-Reyes,J., van Hartesvelt,M. and Koslowsky,D.J. (2010) The impact of mRNA structure on guide RNA targeting in kinetoplastid RNA editing. *PLoS One*, **5**, 1–11.
57. Kuhn,U., Pieler,T., Kühn,U. and Pieler,T. (1996) Xenopus poly(A) binding protein: functional domains in RNA binding and protein-protein interaction. *J.Mol.Biol.*, **256**, 20–30.
58. Imataka,H., Gradi,A. and Sonenberg,N. (1998) A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. *EMBO J.*, **17**, 7480–7489.
59. Cheng,S. and Gallie,D.R. (2007) eIF4G, eIFiso4G, and eIF4B bind the poly(A)-binding protein through overlapping sites within the RNA recognition motif domains. *J. Biol. Chem.*, **282**, 25247–25258.
60. Sladic,R.T., Lagnado,C.A., Bagley,C.J. and Goodall,G.J. (2004) Human PABP binds AU-rich RNA via RNA-binding domains 3 and 4. *Eur.J.Biochem.*, **271**, 450–457.
61. Khanam,T., Muddashetty,R.S., Kahvejian,A., Sonenberg,N. and Brosius,J. (2006) Poly(A)-binding protein binds to A-rich sequences via RNA-binding domains 1+2 and 3+4. *RNA Biol.*, **3**, 170–177.
62. Khacho,M., Mekhail,K., Pilon-Larose,K., Pause,A., Cote,J. and Lee,S. (2008) eEF1A Is a Novel Component of the Mammalian Nuclear Protein Export Machinery. *Mol. Biol. Cell*, **19**, 5296–5308.
63. Kozlov,G., Trempe,J.F., Khaleghpour,K., Kahvejian,A., Ekiel,I. and Gehring,K. (2001) Structure and function of the C-terminal PABC domain of human poly(A)-

binding protein. *Proc. Natl. Acad. Sci. U. S. A.*, **98**, 4409–4413.

64. Kozlov,G. and Gehring,K. (2010) Molecular basis of eRF3 recognition by the MLLE domain of poly(A)-binding protein. *PLoS One*, **5**, 3–8.
65. dos Santos Rodrigues,F.H., Firczuk,H., Breeze,A.L., Cameron,A.D., Walko,M., Wilson,A.J., Zanchin,N.I.T. and McCarthy,J.E.G. (2018) The *Leishmania* PABP1–eIF4E4 interface: a novel 5′–3′ interaction architecture for trans-spliced mRNAs . *Nucleic Acids Res.*, **47**, 1493–1504.
66. Noé,G., Gaudenzi,J. De and Frasch,A.C. (2008) Functionally related transcripts have common RNA motifs for specific RNA-binding proteins in trypanosomes. *BMC Mol. Biol.*, **9**, 1–19.
67. Ledda,M., Di Croce,M., Bedini,B., Wannenes,F., Corvaro,M., Boyl,P.P., Caldarola,S., Loreni,F. and Amaldi,F. (2005) Effect of 3' UTR length on the translational regulation of 5'-terminal oligopyrimidine mRNAs. *Gene*, **344**, 213–220.
68. Hamilton,T.L., Stoneley,M., Spriggs,K.A. and Bushell,M. (2006) TOPs and their regulation. *Biochem. Soc. Trans.*, **34**, 12.
69. Liang,X., Haritan,A. and Uliel,S. (2003) trans and cis Splicing in Trypanosomatids : Mechanism , Factors , and Regulation. *Eukaryot. Cell*, **2**, 830–840.
70. Jensen,B.C., Ramasamy,G., Vasconcelos,E.J.R., Ingolia,N.T., Myler,P.J. and Parsons,M. (2014) Extensive stage-regulation of translation revealed by ribosome profiling of *Trypanosoma brucei*. *BMC Genomics*, **15**, 911.
71. Gupta,S.K., Chikne,V., Eliaz,D., Tkacz,I.D., Naboishchikov,I., Carmi,S., Waldman Ben-Asher,H. and Michaeli,S. (2014) Two splicing factors carrying serine-arginine motifs, TSR1 and TSR1IP, regulate splicing, mRNA stability, and rRNA processing in *Trypanosoma brucei*. *RNA Biol.*, **11**, 715–731.
72. Zinoviev,A., Léger,M., Wagner,G. and Shapira,M. (2011) A novel 4E-interacting protein in *Leishmania* is involved in stage-specific translation pathways. *Nucleic Acids Res.*, **39**, 8404–8415.
73. Costello,J., Castelli,L.M., Rowe,W., Kershaw,C.J., Talavera,D., Mohammad-Qureshi,S.S., Sims,P.F.G., Grant,C.M., Pavitt,G.D., Hubbard,S.J., *et al.* (2015) Global mRNA selection mechanisms for translation initiation. *Genome Biol.*, **16**, 1–21.

74. Fernández-Moya,S.M., García-Pérez,A., Kramer,S., Carrington,M. and Estévez,A.M. (2012) Alterations in DRBD3 Ribonucleoprotein Complexes in Response to Stress in *Trypanosoma brucei*. *PLoS One*, **7**, 1–10.

75. Guerra,N., Vega-Sendino,M., Pérez-Morgado,M.I., Ramos,E., Soto,M., Gonzalez,V.M. and Martín,M.E. (2011) Identification and functional characterization of a poly(A)-binding protein from *Leishmania infantum* (LiPABP). *FEBS Lett.*, **585**, 193–198.

76. Fonseca,B.D., Zakaria,C., Jia,J.J., Graber,T.E., Svitkin,Y., Tahmasebi,S., Healy,D., Hoang,H.D., Jensen,J.M., Diao,I.T., *et al.* (2015) La-related protein 1 (LARP1) represses terminal oligopyrimidine (TOP) mRNA translation downstream of mTOR complex 1 (mTORC1). *J. Biol. Chem.*, **290**, 15996–16020.

77. Philippe,L., Vasseur,J.J., Debart,F. and Thoreen,C.C. (2018) La-related protein 1 (LARP1) repression of TOP mRNA translation is mediated through its cap-binding domain and controlled by an adjacent regulatory region. *Nucleic Acids Res.*, **46**, 1457–1469.

78. Marzluff,W.F., Wagner,E.J. and Duronio,R.J. (2009) Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nat Rev Genet.*, **9**, 843–854.

79. Zhang,J., Tan,D., DeRose,E.F., Perera,L., Dominski,Z., Marzluff,W.F., Tong,L. and Hall,T.M.T. (2014) Molecular mechanisms for the regulation of histone mRNA stem-loop-binding protein by phosphorylation. *Proc. Natl. Acad. Sci.*, **111**, E2937–E2946.

80. Takeuchi,O. and Yamashita,A. (2017) Translational control of mRNAs by 3'-Untranslated region binding proteins. *BMB Rep.*, **50**, 194–200.

81. Singh,A., Minia,I., Droll,D., Fadda,A., Clayton,C. and Erben,E. (2014) Trypanosome MKT1 and the RNA-binding protein ZC3H11: Interactions and potential roles in post-transcriptional regulatory networks. *Nucleic Acids Res.*, **42**, 4652–4668.

82. De Melo Neto,O.P., Standart,N. and De Sa,C.M. (1995) Autoregulation of poly(A)-binding protein synthesis in vitro. *Nucleic Acids Res.*, **23**, 2198–2205.

Table 1. Summary of relevant proteins co-purifying with HA-tagged PABPs*.

Identify	Gene ID	TP-SP[a]		PABP1		PABP2		PABP3	
		1stExp	2ndExp	1stExp	2ndExp	1stExp	2ndExp	1stExp	2ndExp
Translation									
PABP1	LinJ.35.5360	26	25	22	10	4	0	-	0
PABP2	LinJ.35.4200	2	2	1	2	30	17	15	9
PABP3	LinJ.25.0080	-	0	-	0	7	0	9	5
EIF4E4	LinJ.30.0460	3	2	2	0	1	0	-	0
EIF4A1	LinJ.01.0800	2	3	2	8	-	6	-	7
RNA related									
RNA binding protein - RBP23	LinJ.17.0610	3	5	5	0	-	0	-	0
Doubled RNA binding protein - DRBD2	LinJ.35.2240	-	0	-	0	4	5	-	0
Doubled RNA binding protein - DRBD3	LinJ.04.1190	1	-	1	-	-	-	-	-
Nuclear RNA binding domain - NRBD	LinJ.32.0790	-	-	-	-	3	-	1	-
Mitochondrial RNA binding protein - MRP2	LinJ.09.1180	-	-	-	-	2	-	-	-
Zinc Finger and KH domain containing protein - ZC3H41	LinJ.27.1220	5	1	6	2	10	0	4	0
Uncharacterized protein									
Uncharacterized protein containing NTF2 and RRM-like domain	LinJ.18.0300	4	5	1	0	-	0	0	0

*RBPs are shown differentially *co-immunoprecipitated with PABP1 and PABP2/3 in two independent immunoprecipitation (IP) assays.* The values are the hits of proteins bound for each PABP. Only proteins identified with >1 peptide and a probability of >80.0% were considered. [a]TP-SP → phosphorylation mutant of PABP1.

Table 2. Proteins co-purifying with HA-tagged RBP23*.

RBP23/Negative				
Identify	**Gene ID**	**Log_2 (I^a $Ratio^b$)**		**Average >3**
		1st Exp	**2nd Exp**	
Translation				
PABP1	LinJ.35.5360	7.71	7.68	7.69
EIF4E4	LinJ.30.0460	6.05	3.84	4.95
EIF4G3	LinJ.16.1700	5.16	3.06	4.11
RNA related				
RNA binding protein - RBP23	LinJ.17.0610	9.72	8.87	9.29
Zinc Finger and KH domain containing protein - ZC3H41	LinJ.27.1220	4.7	3.69	4.19
Pentatricopeptide repeat protein – PPR	LinJ.21.1510	4.85	3.1	3.98
CCR4-NOT transcription complex subunit - NOT1	LinJ.21.0880	5.07	1.38	3.22
CCR4-NOT transcription complex subunit - CAF1	LinJ.22.1480	4.01	2.42	3.21
Uncharacterized protein				
Uncharacterized protein containing NTF2 and RRM-like domain	LinJ.18.0300	5.87	4.91	5.39
Uncharacterized protein	LinJ.04.0140	4.84	2.83	3.83
Uncharacterized protein	LinJ.19.0020	4.4	3.21	3.81
Uncharacterized protein	LinJ.05.0450	3.56	3.37	3.46

*11 polypeptides are shown which specifically co-immunoprecipitated with the HA-tagged RBP23 in two independent immunoprecipitation (IP) experiments, with a minimum of 8-fold increase over the negative control (Log_2 (I Ratio) >3).

[a]I= Intensity

[b]R = Ratio between the intensity observed for the IP with extracts derived from the cells expressing HA tagged RBP23 divided by the intensity detected for IP with the extracts from wild type cells.

Table 3. Proteins co-purifying with HA-tagged DRBD2*.

DRBD2/Negative				
Identify	**Gene ID**	**Log_2 (I^a Ratio^b)**		**Average >3**
		1stExp	**2ndExp**	
Translation				
PABP3	LinJ.25.0080	4.11	4.17	4.14
PABP2	LinJ.35.4200	4.51	3.63	4.07
EIF4G3	LinJ.16.1700	3.75	3.14	3.44
RNA related				
Zinc finger domain containing protein - ZC3H39	LinJ.19.0290	8.81	8.26	8.54
Zinc finger domain containing protein - ZC3H40	LinJ.19.0300	8.53	8.29	8.41
Doubled RNA binding protein - DRBD2	LinJ.35.2240	7.59	7.36	7.47
Mitochondrial edited mRNA stability factor 1 subunit	LinJ.28.0040	7.78	7.16	7.47
Guide RNA associated protein - GAP2	LinJ.22.0520	6.92	5.03	5.98
Kinetoplast polyadenylation factor – KPAF1	LinJ.18.0010	5.54	6.36	5.95
RNA binding protein – RBP43	LinJ.25.0290	5.58	5.53	5.56
Mitochondrial RNA binding protein – MRB11870	LinJ.33.1320	6.01	4.94	5.48
Guide RNA associated protein - GAP1	LinJ.33.2870	5.69	5.23	5.46
Pumilio protein - PUF6	LinJ.33.1210	4.78	5.51	5.14
Mitochondrial RNA binding protein – MRB10130	LinJ.36.5000	5.32	4.87	5.1
Guide RNA binding protein - gBP21	LinJ.27.0980	5.08	4.9	4.99
Polyadenylation mediator complex protein - PAMC1	LinJ.20.0960	5.43	3.93	4.68
Mitochondrial RNA binding protein - MRP2	LinJ.09.1180	4.32	4.69	4.51
Zinc finger domain containing protein - ZC3H31	LinJ.36.0800	3.59	4.88	4.24
RNA binding protein	LinJ.21.0490	4.15	4.32	4.24
Mitochondrial RNA binding protein - MRB4160	LinJ.31.0670	3.11	5.28	4.2
RNA-binding protein - RBP12	LinJ.18.0590	3.72	4.3	4.01
'Cold-shock' DNA binding domain containing protein	LinJ.14.1210	2.43	5.58	4
Three RRMs domain containing protein - TRRM2	LinJ.30.1170	3.06	4.27	3.66
Guide RNA binding protein	LinJ.08.1080	3.93	2.92	3.43
Pumilio protein - PUF1	LinJ.36.0050	3.58	3.22	3.4
Three RRMs domain containing protein - TRRM1	LinJ.27.2020	3.01	3.18	3.1
Catalytic activity				
ATP-dependent 6-phosphofructokinase – PFK	LinJ.29.2620	2.51	5.04	3.77
ATP-dependent RNA helicase – HEL67	LinJ.32.0410	3.56	3.72	3.64
Nucleolar RNA helicase II	LinJ.05.0140	3.64	3.61	3.62
Uncharacterized protein				
Uncharacterized protein	LinJ.19.0020	5.67	5.63	5.65
Uncharacterized protein	LinJ.20.0090	3.32	3.24	3.28

*30 polypeptides are shown which specifically co-immunoprecipitated with the HA-tagged DRBD2 in two independent immunoprecipitation (IP) experiments, with a minimum of 8-fold increase over the negative control (Log_2 (I Ratio) >3).

[a]I= Intensity

[b]R = Ratio between the intensity observed for the IP with extracts derived from the cells expressing HA-tagged DRBD2 divided by the intensity detected for IP with the extracts from wild type cells.

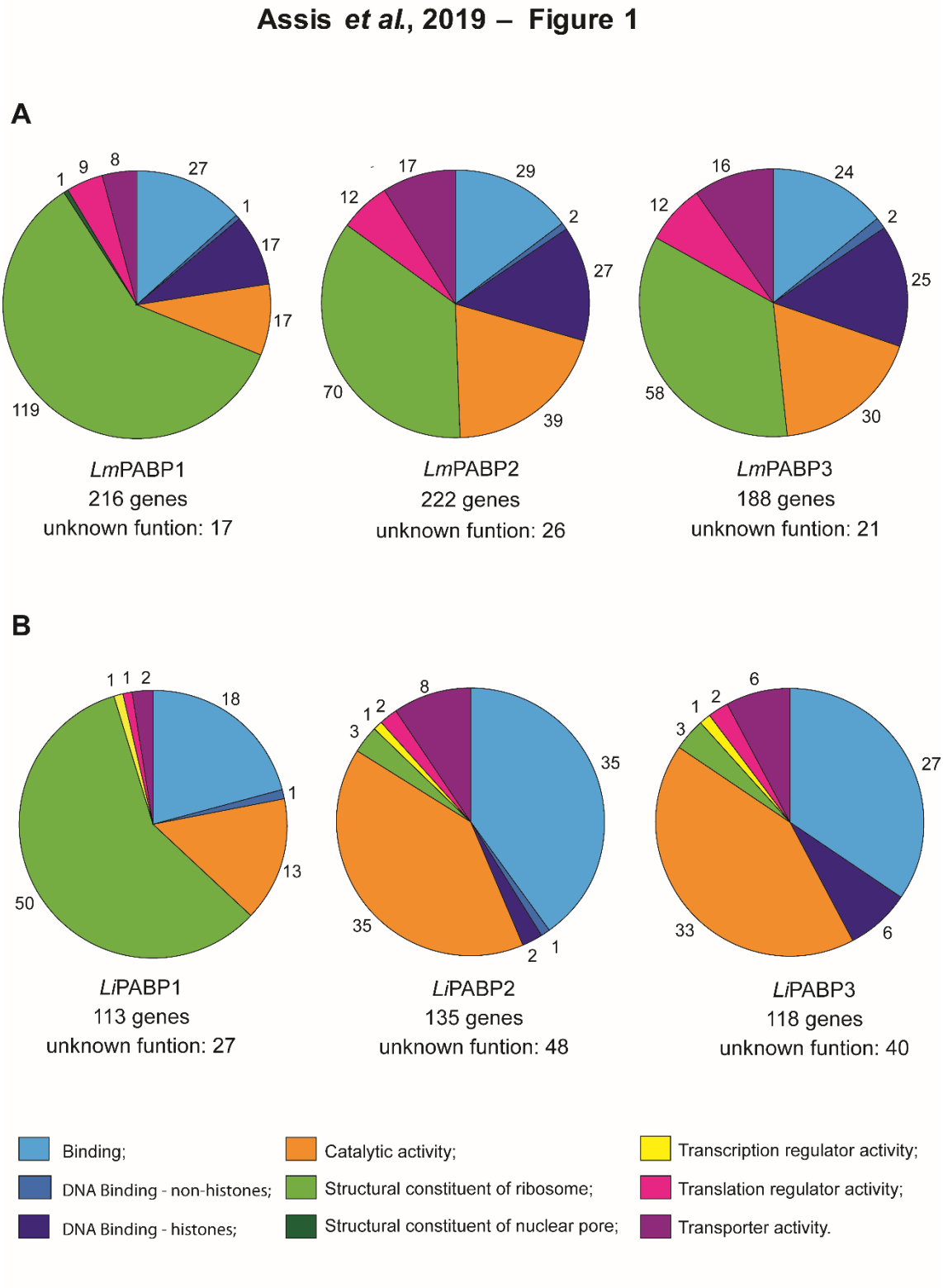


Figure 1. Analysis of mRNAs populations associated with the three PABPs in *Leishmania sp.* Upregulated genes in at least two of three available RNA-seq datasets were manually classified and grouped using the gene ontology terms according their molecular function. 'Upregulated' means at least twofold more abundant than in the negative control; **A)** mRNAs groups associated to native PABP1, PABP2 and PABP3 from *L. major* (SOLiD)*;* **B)** mRNAs groups associated to recombinant PABP1, PABP2 and PABP3 from *L. infantum* (Illumina)*.*

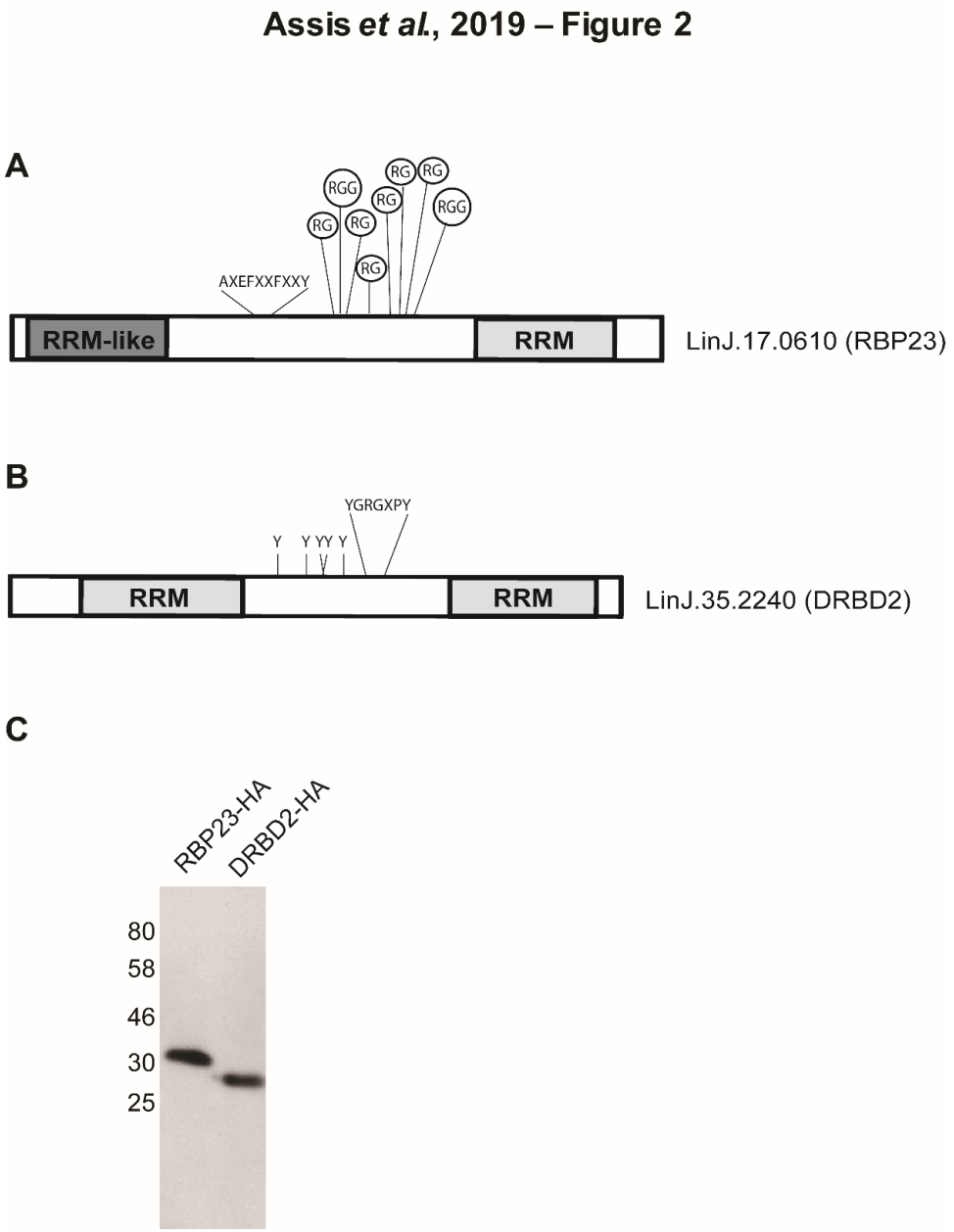


Figure 2. Expression of proteins HA-tagged by immunodetection with a monoclonal anti-HA. A) Schematic representation of *L. infantum* RBP23. The protein presents a RRM-like in N-terminal and a typical RRM domain in C-terminal. Eight putative arginine methylation sites are indicated (RG or RGG) as a conserved set of aromatic amino acids between both RRMs. **B)** Schematic representations of *L. infantum* DRBD2. Besides two well-defined RRMs, the protein presents a highly conserved sequence of amino acid and several tyrosines; **C)** Western-blot assays of cell extract expressing RBP23-HA and DRBD2-HA with the size approximate of 32 and 30 kDa, respectively.

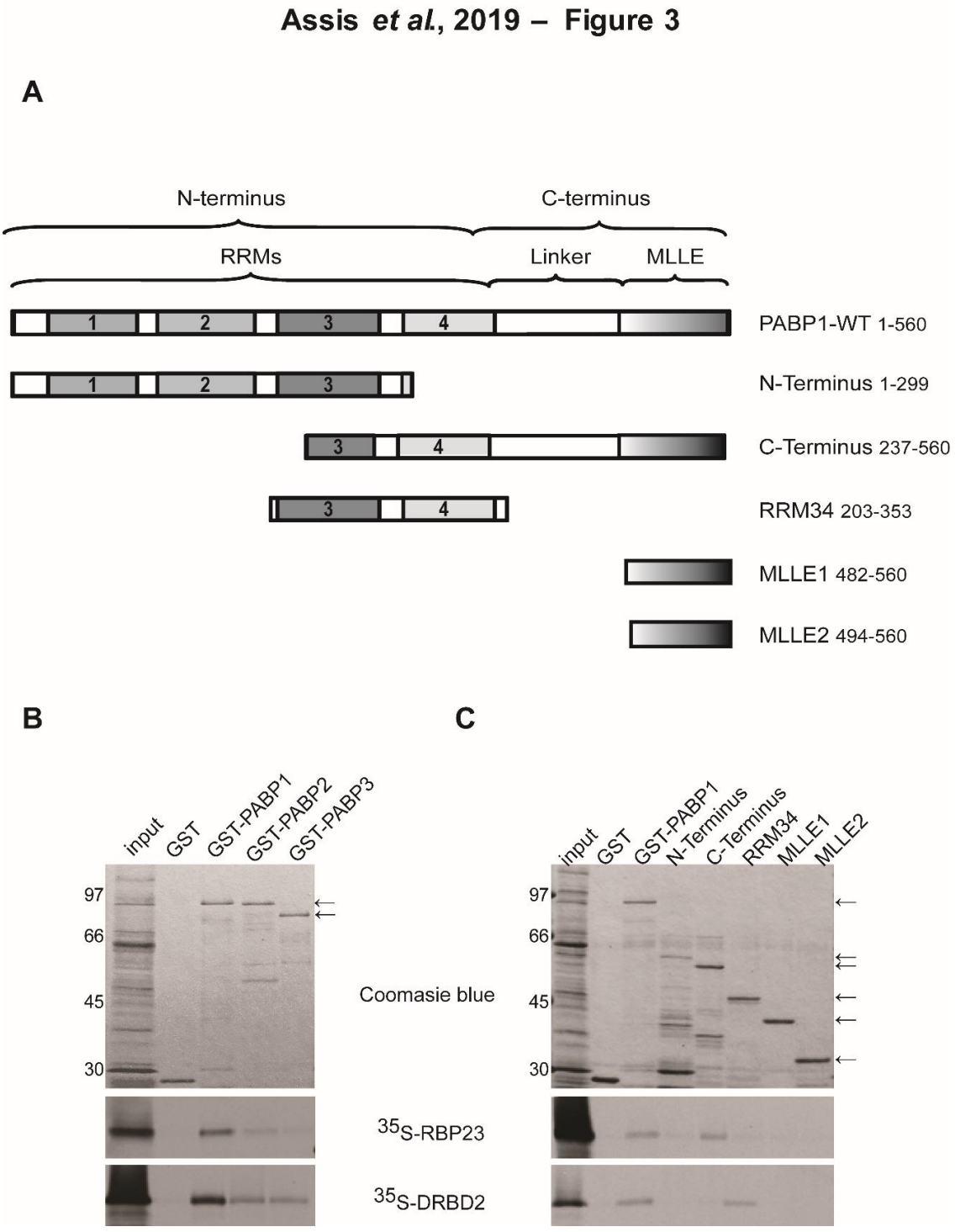


Figure 3. **Interactions between three PABPs and PABP1 variants with RNA binding proteins RBP23 and DRBD2 from *L. infantum*. A)** Schematic representation of *L. infantum* PABP1 homologue and variants: PABP1-WT (wild type); N-terminus; C-terminus; RRM34, MLLE1 and MLLE2; **B)** Pull down assays between radiolabeled RBP23 or DRBD2 and three recombinant PABPs; **C)** Pull down assays between radiolabeled RBP23 or DRBD2 and recombinant PABP1 truncations; The upper panels shows the Coomassie-blue stained gel indicating the radiolabeled protein input, the GST recombinant (negative control) and the GST-PABP homologues or the GST-PABP1 truncations; The panels below are autoradiographs showing the result of the interactions between radiolabeled RBP23 or DRBD2 and recombinant proteins. The recombinant proteins are indicated by arrows and the sizes of molecular weight markers are shown on the left.

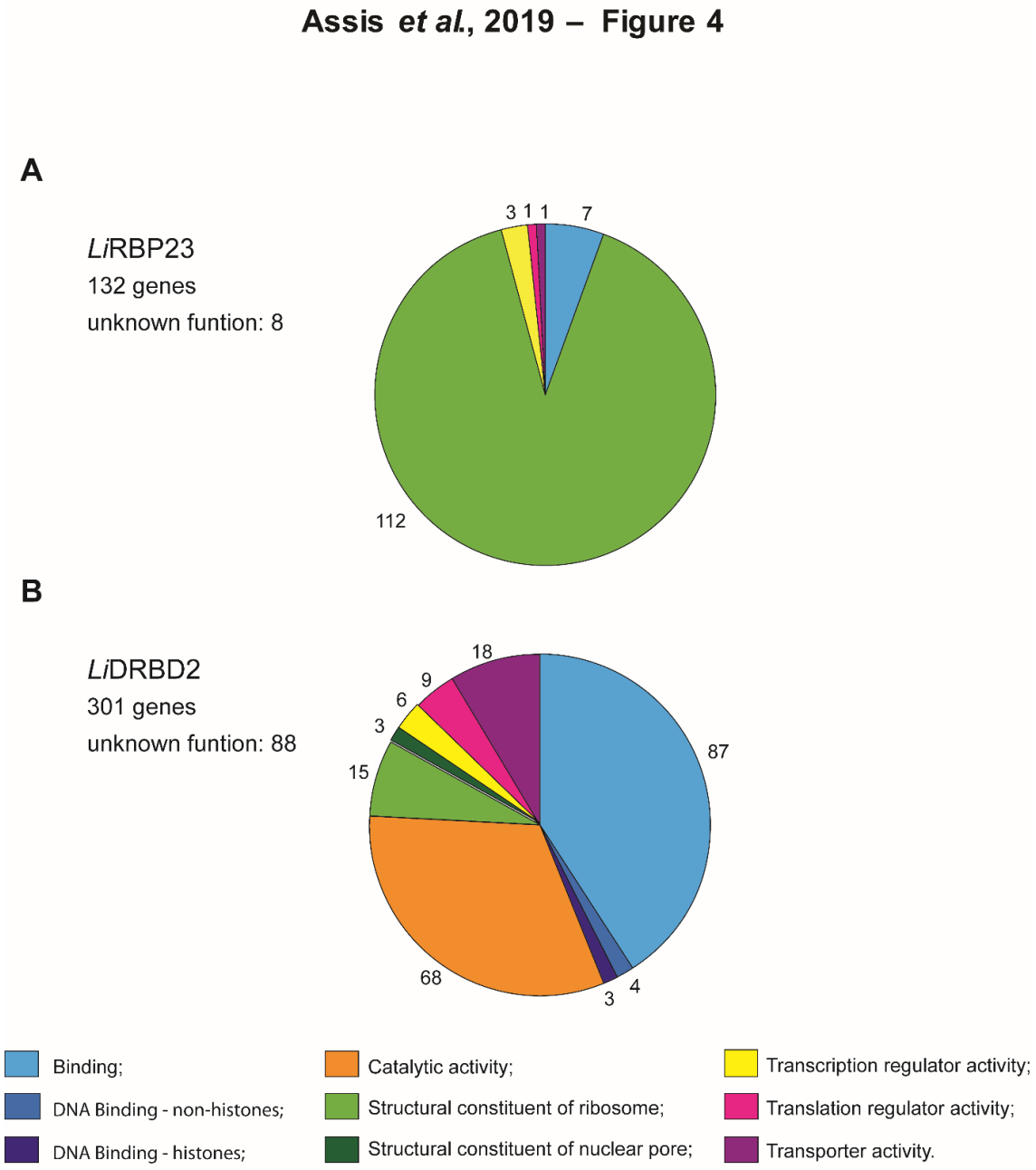


Figure 4. Analysis of mRNAs populations associated with RBP23 and DRBD2 in *L. infantum*. Genes were classified and grouped as detailed in Figure 1; **A)** Upregulated mRNAs groups associated to RBP23. **B)** Upregulated mRNAs groups associated to DRBD2.

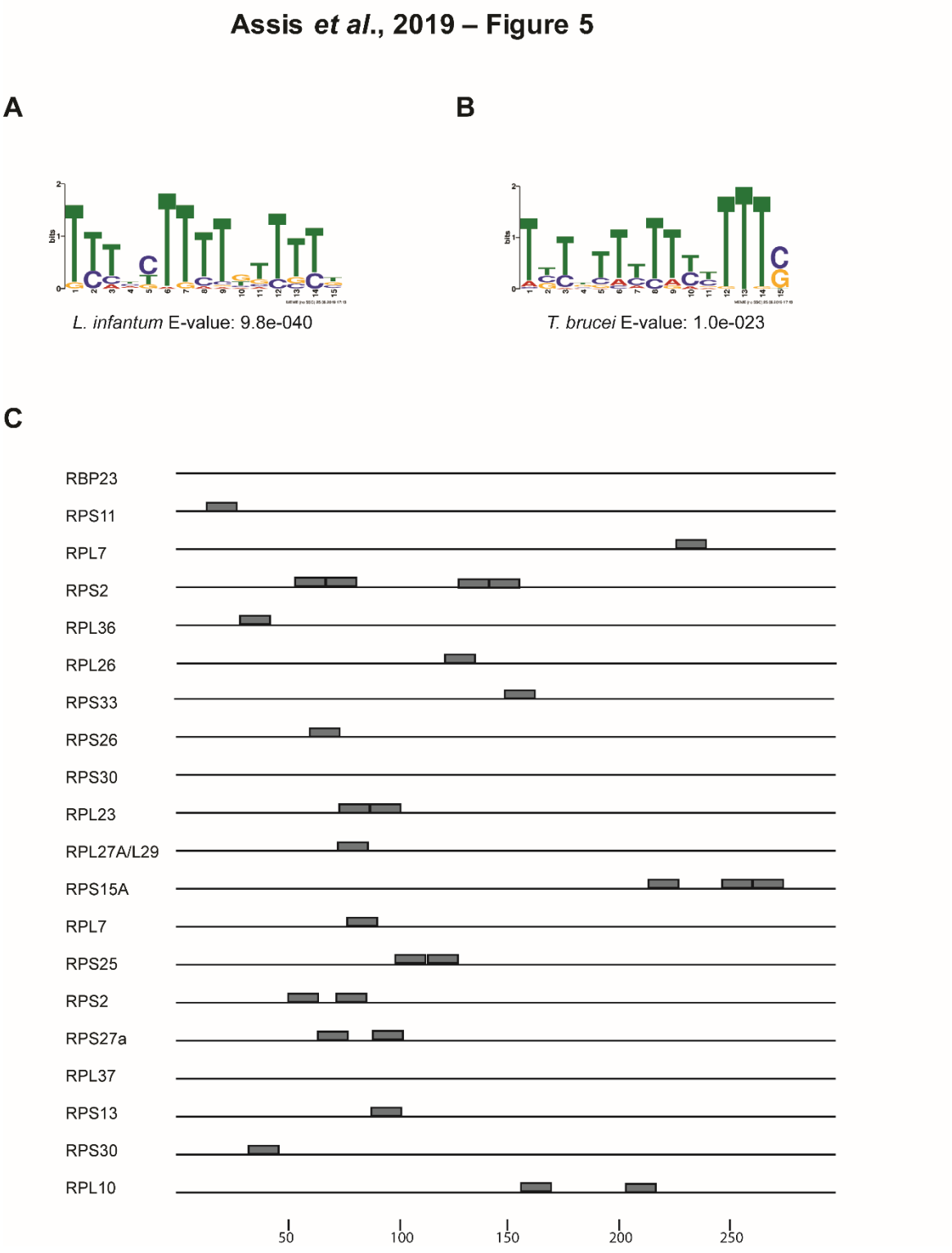


Figure 5. Sequence elements in 3' UTR of trypanosomatid ribosomal protein mRNAs. A) Putative upstream motif for *L. infantum* RP mRNAs; **B)** Putative upstream motif for *T. brucei* RP mRNAs. Both logos were generated by MEME (Multiple Em for Motif Elicitation) and the relative height of each nucleotide represents the conservation measure at that specific position**; C)** Localization of T-rich motifs in the 3'UTR of the 20 mRNAs more enrichment associated with RBP23 of *L. infantum*.

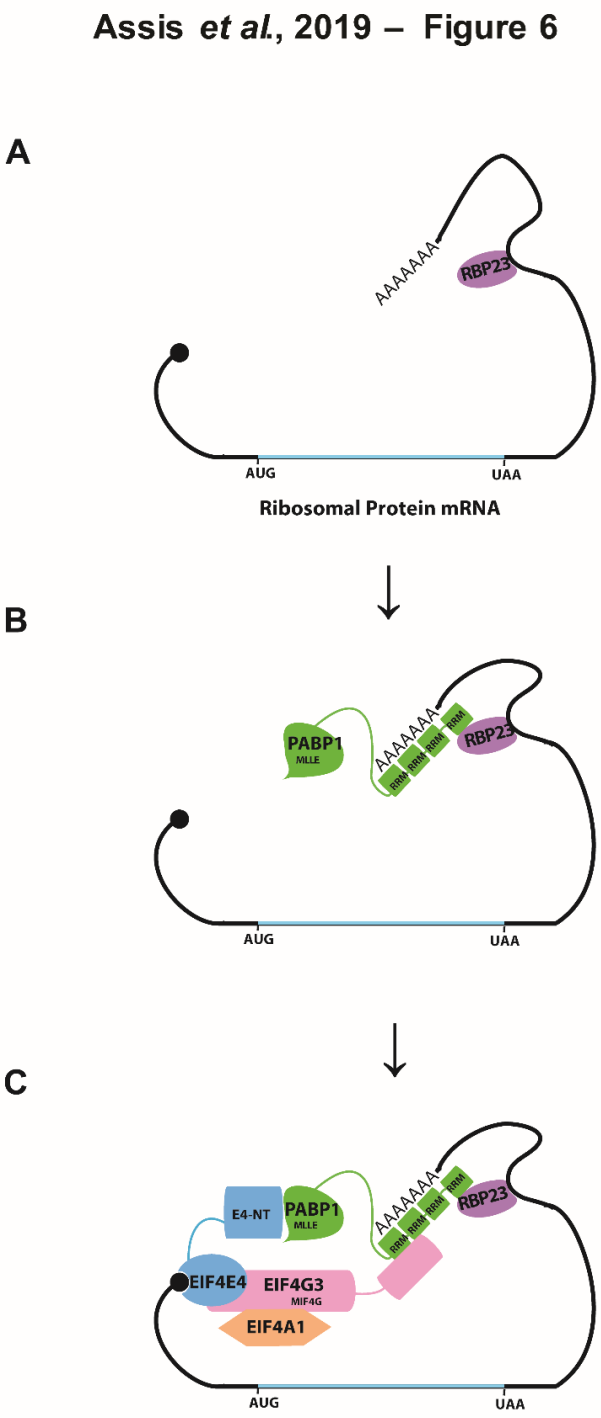
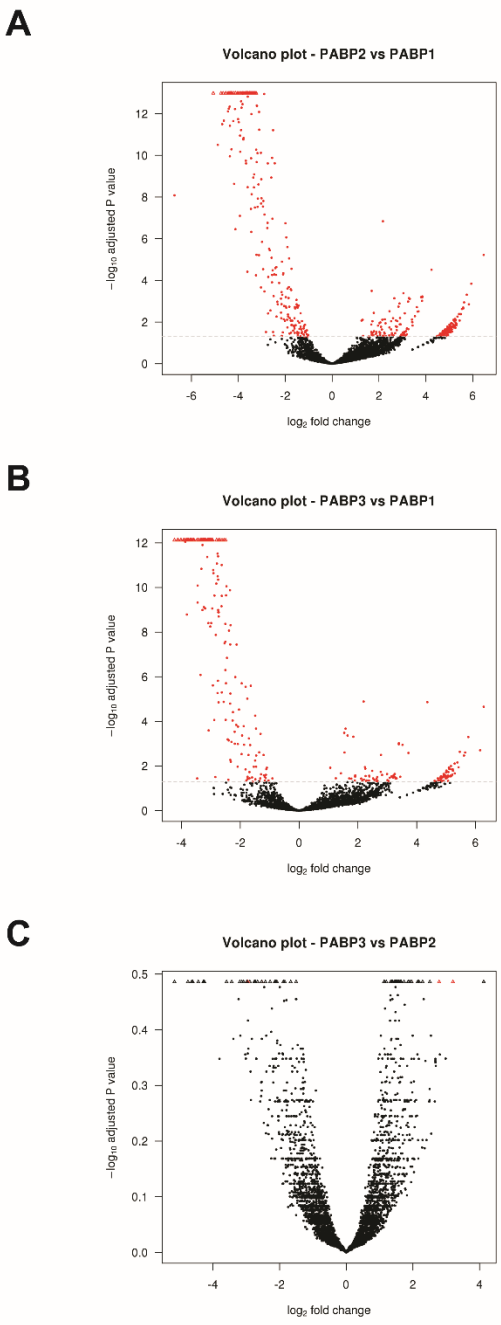


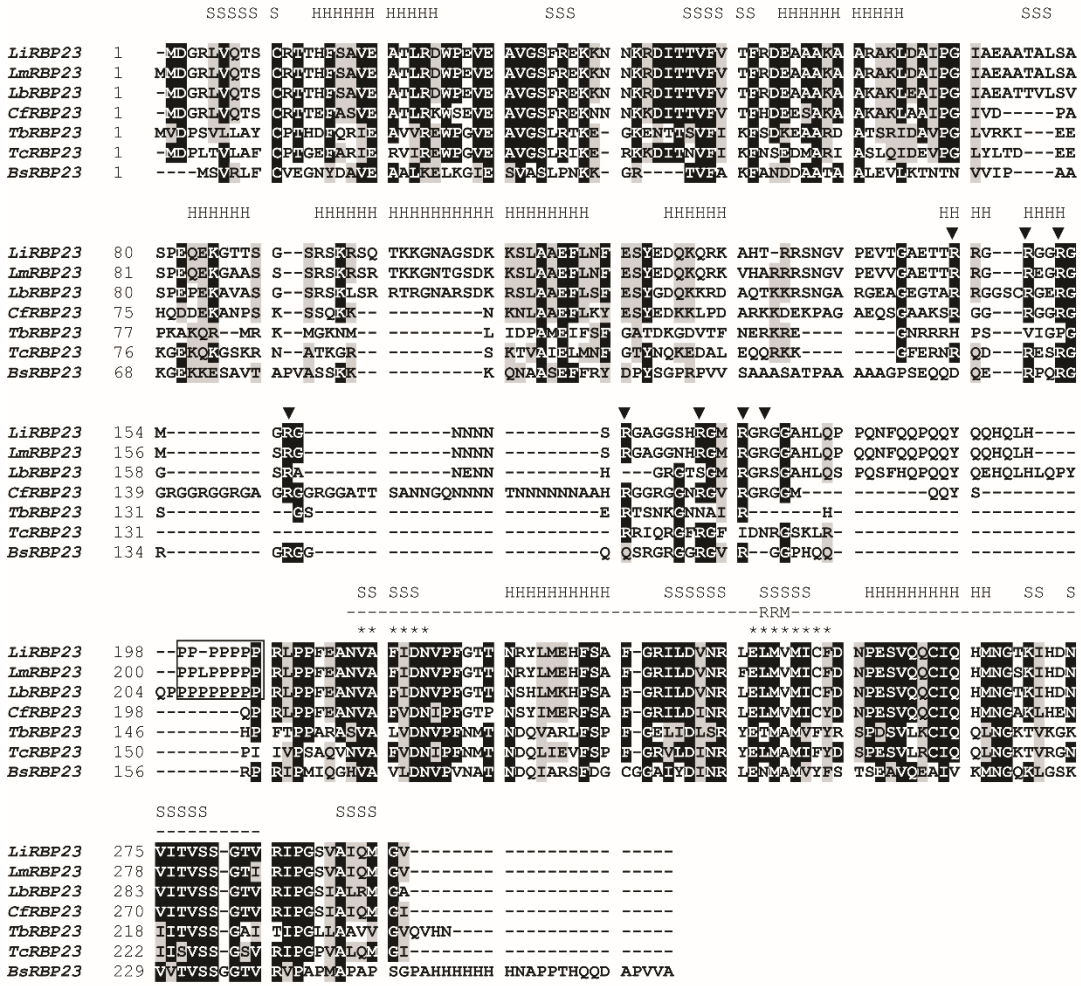
Figure 6. Proposed model for RBP23 selection mRNA to translation. (A) RBP23 interacts with 3' UTR motif on the ribosomal protein mRNA, (B) then it interact with PABP1 probably via RRM4. (C) This set of proteins recruits the complex EIF4E4/EIF4G3/EIF4AI.

Assis *et al.*, 2019 – Supplementary Figure 1



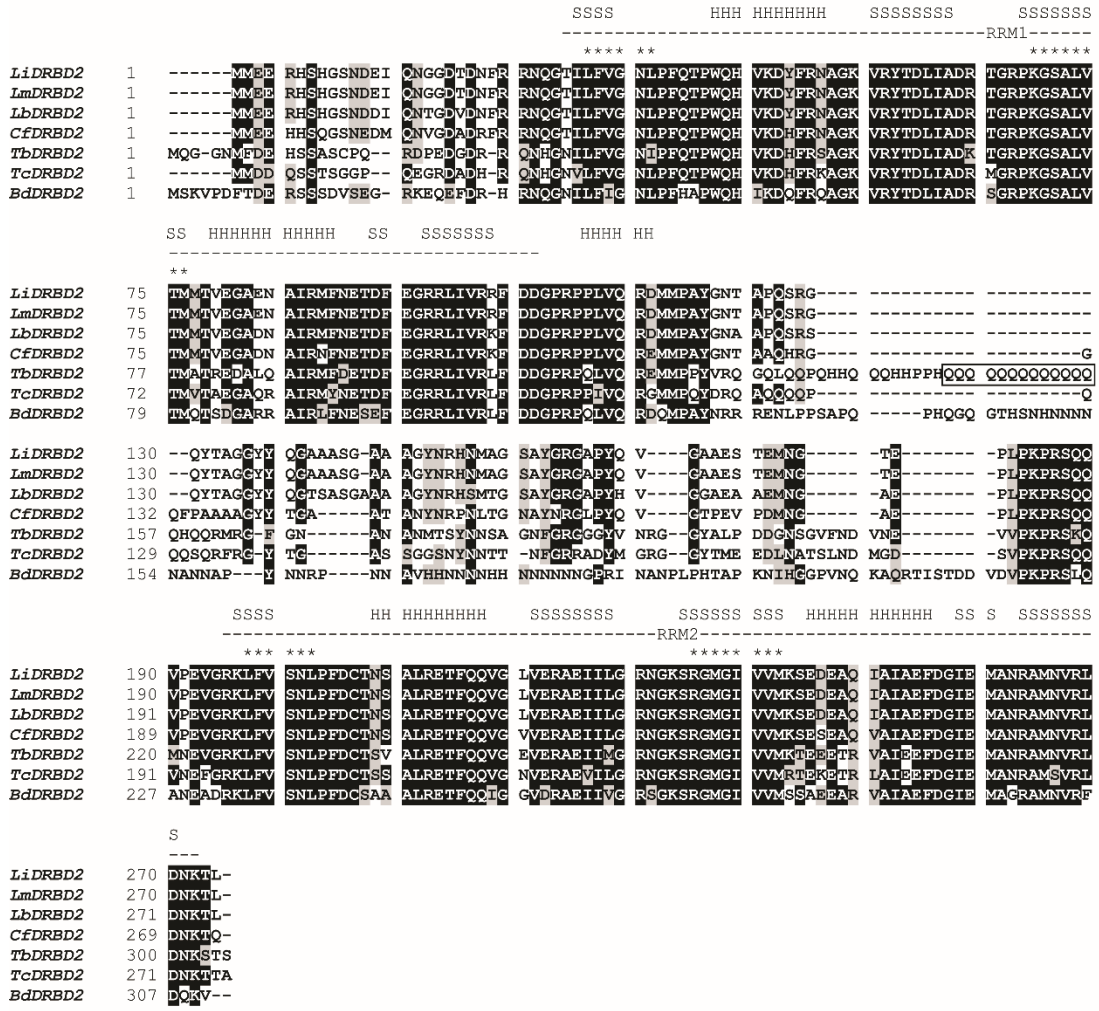
Supplementary Figure 1. Volcano Plot comparing mRNA content of *L. infantum* PABPs RNA-seq dataset. Axis Y represent the statistical test significance 0.05 (>0.05 red and <0.05 black) and the axis X represent the genes at least 4-fold up- or down-regulated (log_2 fold change >2); Δ very high value to be ranked; **A)** PABP2xPABP1 **B)** PABP3xPABP1 **C)** PABP3xPABP2.

Assis *et al.*, 2019 – Supplementary Figure 2



Supplementary Figure 2. Sequences analysis of the RBP23 from seven organisms of the kinetoplastida class. MAAFT alignment comparing the sequences showing the conserved and divergent motifs in seven organisms of the kinetoplastida class: *L. infantum*, *L. major*, *L. braziliensis*, *C. fasciculata*, *T. brucei*, *T. cruzi*, *B. saltans.* Identical amino acids in more than 60% of the sequences are highlighted in dark gray, while amino acids defined as similar, based on the BLOSUM60 matrix, on more than 60% of the sequences are highlighted in light gray. When necessary, spaces were inserted within the various sequences (dashes) to allow better alignment. The RRMs of *L. infantum* sequence are indicated as well as the secondary structures predicted. H and S represent α-helices and β-sheet, respectively. The asterisks define the two RNP motifs. ▼ define the putative arginine methylation sites (RG or RGG). The P-repeat sequence of *Leishmania* genus was highlighted. Access numbers: *LiRBP23* - LinJ.17.0610; *LmRBP23* - LmjF.17.0550; *LbRBP23* - LbrM.17.0540; *CfRBP23* - CFAC1_280015000; *TbRBP23* - Tb927.10.11270; *TcRBP23* - TcCLB.507711.40; *BsRBP23* - CUG90143.1.

Assis *et al.*, 2019 – Supplementary Figure 3



Supplementary Figure 3. Sequences analysis of the DRBD2 from seven organisms of the kinetoplastida class. The alignment was carried out as described for Figure 1. The Q-repeat sequence of *T. brucei* was highlighted. Relevant access numbers: *LiDRBD2* - LinJ.35.2240; *LmDRBD2* - LmjF.352200; *LbDRBD2* - LbrM.34.2130; *CfDRBD2* - CFAC1_300070700; *TbDRBD2* - Tb927.9.13990; *TcDRBD2* - TcCLB.510755.120 and *BsDRBD2* - CUG88059.1.

4 ARTIGO 2

4.1 Caracterização da ZC3H41, contendo domínio dedo de zinco do tipo CCHC, parceiro proteico dos homólogos das proteínas de ligação à cauda poli-A (PABP) em *Leishmania infantum*

Artigo anexo na página seguinte.

A doutoranda e co-autora desse artigo em processo de finalização experimental foi responsável pela escrita, pela experimentação e análises do sequenciamento dos mRNAs associados a ZC3H41. A doutoranda co-orientou as alunas Irassandra de Aquino e Yallen Santos que realizaram amplificação do gene *ZC3H41*, clonagem em vetores de expressão, obtenção de anticorpos, curvas de crescimento, imunoprecipitações, *western-blot* e espectrometria de massas. Resumidamente, a ZC3H41 co-precipita não apenas com a PABP1, mas também com PABP2 e PABP3. A análise de RNAseq revelaram que a ZC3H41 se liga preferencialmente aos mRNAs de proteínas ribossomais, semelhante aos parceiros PABP1 e RBP23. Para ensaios de localização subcelular, o soro policlonal foi produzido contra ZC3H41 e com sua especificidade confirmada.

Characterization of ZC3H41, a Zinc finger CCHC-type protein partner of Poly(A) Binding Protein (PABP) homologues from *Leishmania infantum*

Irassandra R. P. U. C. de Aquino[1], Ludmila A. Assis[1,2], Yallen S. de Melo[1], Tamara D. C. da Costa Lima[3], Maria J. R. Bezerra[1,2], Antônio M. Rezende[1] and Osvaldo P. de Melo Neto*[1]

[1]Department of Microbiology, Aggeu Magalhães Institute – Oswaldo Cruz Foundation, Recife, PE, Brazil
[2] Federal University of Pernambuco, Recife, PE, Brazil
[3] University Center Tabosa de Almeida, Caruaru, PE, Brazil.

*Corresponding author e-mail: opmn@cpqam.fiocruz.br

Keywords: Trypanosomatidae. ZC3H41. PABP1

ABSTRACT

Translation initiation in *Leishmania* is characterized by the presence of multiple subunits for the heterotrimeric eIF4F initiation complex, forming a minimum of five eIF4Fs complexes, as well as three Poly (A) Binding Protein (PABP) homologues. These presumably bind to different RNA binding proteins (RBPs) as well as different mRNA targets. Here we report the characterization of ZC3H41, a Zinc Finger CCHC-type protein, which was first found to be associated to PABP1. *L. infantum* transgenic cell lines were generated expressing HA-tagged ZC3H41, with the ectopic protein showing a constant expression during a growth curve. Cytoplasmic extracts of these cells were used in immunoprecipitation, and proteins and mRNAs bound to ZC3H41 were analyzed by mass spectrometry and RNAseq, respectively. ZC3H41 was found to co-precipitate not only with PABP1, but also with PABP2 and PABP3. The ZC3H41 RNAseq analysis revealed it to preferably bind to ribosomal protein mRNAs, similar to another PABP1 partner, RBP23. Polyclonal serum was produced against ZC3H41 and its specificity confirmed as the native

protein expression. These experiments shed light on the function of the *Leishmania* ZC3H41 and its role in mRNA translation.

INTRODUCTION

In higher eukaryotes, translation initiation starts when the smaller ribosomal subunit, the initiator tRNA (tRNAi) and associated translation initiation factors (eIFs) form the pre-initiation complex (PIC) which then associates with the mRNA at its 5' end. Once attached, PIC generally identifies the first AUG codon and then recruits the larger ribosomal subunit to start the mRNA translation. Another translation factor, eIF4F, a complex of three different polypeptides (eIF4A, eIF4E and eIF4G), plays an essential role in the binding of the mRNA to PIC. Within eIF4F, eIF4E recognizes the mRNA by binding to the cap structure at its 5' end. eIF4A is an RNA-dependent ATPase and RNA helicase that removes secondary mRNA structures, facilitating the scanning of the mRNA 5'untranslated region (5'UTR) and the identification of the AUG codon. The largest eIF4F subunit, eIF4G, is a multi-domain protein with multiple binding sites for other polypeptides, such as eIF4E, eIF4A, eIF3 (part of PIC and directly bound to the smaller ribosomal subunit) and the Poly (A) Binding Protein (PABP). PABP binds to the mRNAs' 3' Poly (A) tail and its interaction with eIF4G allows their circularization, stimulating translation and the recycling of ribosomes (1–4).

The best known trypanosomatids, parasitic protozoans which diverged early from the main eukaryotic lineage, are responsible for diseases of worldwide impact and medical interest, such as the *Leishmania* species (leishmaniasis) and the *Trypanosoma brucei* (sleep disease). These organisms have peculiar molecular characteristics, such as the post-transcriptional regulation of their gene expression, mediated mainly by the control of mRNA stability and translation (5). Several homologues of eIF4F subunits and PABPs have been characterized in trypanosomatids and two distinct eIF4F-like complexes, based on different eIF4E and eIF4G homologues, have been better described. The complex formed by the interaction between EIF4E4/EIF4G3 has been directly implicated in translation, whereas the evidence for a role in translation for EIF4E3/EIF4G4 is less clear. A specific interaction between PABP1 and EIF4E4/EIF4G3 has been recently shown,

while PABP2 and PABP3 (not present in *T. brucei*) most likely interacts with EIF4E3/EIF4G4 and perform other functions in the nucleus (6–12).

Regulation of translation initiation may occur through several mechanisms affecting the cap structure, the eIF4G-eIF4E interaction, the recruitment of the 43S initiation complex, the 80S assembly and the elongation of the poly-A tail, all of which might require the action of RNA binding proteins (RBPs) which bind to the mRNAs (13). RBPs are *trans*-acting elements capable of interacting with regulatory motifs on the mature transcripts, the *cis*-acting elements. The majority of *cis*-acting elements have been detected within the 3' untranslated region (UTRs) of the mRNAs and the *trans*-acting factors are classified according to the types of structural domains involved in RNA-binding: RRMs domains, CCCH zinc finger motifs and PUF domains (14, 15). The RBPs seem to be involved during all stages of the trypanosomatids' gene expression pathways, such as mRNA processing, transport, decay and translation (16, 17).

The CCCH-type zinc finger proteins are defined by the Cys–X_{4-15}–Cys–X_{4-6}–Cys–X_3–His motif and bind almost exclusively to single-stranded RNAs with a preference for AU-rich RNA elements (18, 19). These proteins are usually related to mRNA decay processes and their regulation, however there are also zinc finger domain proteins involved during translation initiation. For instance, the β-subunit of the translation initiation factor eIF2 (eIF2β), which has at its C-terminus ZF-like domains, participates in mRNA binding and also during recognition of the translation initiation codon (20).

In previous mass spectrometry analysis, a protein containing a CCCH-type zinc finger domain, a KH domain and a conserved C-terminal helicase domain was identified co-precipitated with PABP1 along with several other RBPs (21). Its *T. brucei* homologue, named ZC3H41, is a member of the zinc finger protein family and when compared with *L. major* orthologue has an identity of 48% (22)**.** This protein is associated to ZC3H11 and MKT1, both involved in mRNA stabilization since their binding to AU-rich elements in 3' UTRs regulate mRNA degradation (14, 23). Here, we focused on the characterization of the *Leishmania infantum* ZC3H41. This was achieved through mass spectrometry and RNAseq of ZC3H41 bound protein partners and mRNA targets, and through a comparison with similar assays carried

out with PABP1 and related RBPs. These helped define ZC3H41 as a component of distinct mRNA binding complexes.

MATERIALS AND METHODS

PCR and cloning

The *Leishmania infantum* genomic DNA from the MHOM/MA/67/ITMAP-263 strain was isolated using DNAzol (Life Technologies). The full length *ZC3H41* gene was amplified using primers flanked by sites for the restriction enzymes BamHI and HindIII (5' primer **GGATCC**ATGTCTCAACAGAATGCTGTCGAG and 3' primer **AAGCTT**GCGACGCTTGCCGAAGTTCTG)**.** PCR fragments were first cloned into the pGEM-T Easy vector (Promega) followed by sequencing and subcloning. For immunoprecipitation assays, the *ZC3H41* gene was subcloned into BamHI-HindIII sites of a modified version of the *Leishmania* expression vector pSPBT1YNEOα (24). The modification consisted of a 27 nucleotides extension encoding the HA epitope (YPYDVPDYA), added immediately after the coding sequence (the HindIII site) and before the translation stop codon. To generate the polyclonal serum, the gene were subcloned into the same sites of pET-21a (Novagen) and pGEX3T4 (GE Healthcare) for the expression of recombinant His- and GST-tagged ZC3H41, respectively.

Parasite growth and expression analysis

Leishmania infantum MHOM/MA/67/ITMAP-263 promastigotes were cultured in Schneider's insect medium supplemented with 10% heat-inactivated Fetal Bovine Serum and 2% hemin at pH 7.2, 25ºC. For the expression analysis, late exponentially grown *L. infantum* were harvested and resuspended directly into SDS-PAGE sample buffer (10% SDS, 1 M Tris-HCl pH 6.8, 50% glycerol, bromophenol blue and 5% 2-β-mercaptoethanol). For the growth curves, stationary cells were diluted to a pre-established cell density of 10^6 cells/ml. At selected time points (0h, 6h, 24h, 48h, 72h, 96h, and 120h), samples were harvested and the sediment was resuspended in SDS-PAGE sample buffer. The cell extracts were submitted to denaturing SDS-PAGE, transferred to a membrane and then blotted with mouse monoclonal antibodies directed against the HA epitope (Anti-HA monoclonal antibody, 100 ηg/ml, Applied Biological Materials). Transfection procedures used for circular plasmids

(episomal expression) were carried out by electroporation. Briefly, grown cells were harvested and washed once in HEPES-NaCl buffer (21 mM HEPES pH 7.05, 137 mM sodium chloride, 5 mM potassium chloride, 0.7 mM disodium phosphate, 6 mM glucose) and then transferred to cuvettes containing the plasmid DNA. Electroporation was carried out with the Gene Pulser Xcell™ electroporation system (Bio-Rad) with a pulse of 450 V, 500 µF. Cells transfected with the *pSPBT1YNEOα* constructs were selected with G418 (20 µg/ml, Sigma).

Expression and purification of recombinant proteins

For the expression of either His- or GST-tagged recombinant proteins, plasmids were transformed into *Escherichia coli* BL21 cells. The transformed bacteria were grown in LB medium with ampicillin (100 µl/mL) and induced with IPTG (isopropyl-D-thiogalactopyranoside). The induced cells were sedimented, resuspended in PBS (Phosphate Buffered Saline), and lysed by sonication (6 pulses with 60s duration, 30s intervals and 30% amplitude) with the Vibra-Cell™ Ultrasonic Liquid Processor VCX 500 / VCX 750 (SONICs®). Protein purification was performed as previously described (25) with Ni-NTA agarose (Qiagen) or glutathione-4B-Sepharose (Amersham Biosciences). Proteins were analyzed in 15% SDS-PAGE stained with Coomassie blue R-250. For the quantification of the recombinant proteins, serial dilutions were compared in Coomassie-stained gels with serial dilutions of known concentrations of bovine serum albumin (BSA). His-tagged protein was used for immunization of rabbits whereas the GST-tagged protein was used to validate the specificity of the antibodies obtained.

Antibody production

Rabbit antisera was raised against ZC3H41 by immunizing adult New Zealand White rabbits with the respective recombinant His-tagged protein. To reduce the background, the antibodies were affinity purified and validated in western blot assays against different concentrations of total extract of *L. infantum* and recombinant GST-tagged protein. The western blotting was performed with the Immobilon-P polyvinylidene difluoride (PVDF) membrane (Millipore), using the purified primary anti-ZC3H41 antibody diluted at 1:1500 and as the second antibody peroxidase-conjugated goat anti-rabbit IgG serum (Jackson ImmunoResearch Laboratories)

diluted at 1:10000. Reactions were detected by enhanced chemiluminescence.

Extract preparation for immunoprecipitation

For mass spectrometry and RNA sequencing analysis, total cytoplasmic extracts from wild type *L. infantum* and recombinant strains expressing HA-tagged ZC3H41, were obtained after lysing the cells through cavitation. First, late exponentially grown *L. infantum* promastigotes were harvested and washed once in ice cold PBS, followed by resuspension in HEPES-lysis buffer (20 mM HEPES-KOH, pH7.4, 75 mM potassium acetate, 4 mM magnesium acetate, 2 mM DTT, supplemented with and EDTA-free protease inhibitors from Roche) to a concentration of 1-2x109 cells/ml. Lysis was carried out using nitrogen cavitation as described (26) but with modifications. Briefly, the resuspended cells were transferred into the cavitation chamber of the cell disruption vessel (Parr Instruments) and incubated at 4ºC under 70 bar pressures for 40 minutes, followed by rapid decompression and lysis. The lysates were submitted to centrifugation at 17,000g for 10 minutes to remove cellular debris and the supernatants, the cytoplasmic extracts, aliquoted and stored at -80ºC.

Immunoprecipitation

For the immunoprecipitation assays (IPs), cytoplasmic extracts of *L. infantum* wild type and HA-tagged ZC3H41, were mixed with Pierce™ Anti-HA Magnetic Beads as per manufacturer's instructions. Briefly, 0.2 mg of the anti-HA magnetic beads were washed three times with PBS followed by the incubation with the cytoplasmic extract for 1 h at 4ºC. After, the depleted supernatant was removed and the beads washed three times with PBS, the resulting, specifically bound, immunoprecipitated antigen-antibody complexes were eluted in SDS-PAGE sample buffer. They were then assayed by 15% SDS-PAGE and western blotting using antibodies against the HA tag diluted 1:3000 to confirm the efficiency of the precipitation reaction. IPs were performed in two or three independent experiments for mass spectrometry and RNAseq, respectively.

Mass spectrometry analysis

Eluted proteins from IP of wild type *L. infantum* and HA-tagged ZC3H41 were loaded onto 10% SDS-PAGE gels and allowed to migrate into the resolving gel, when the electrophoresis was interrupted prior to protein fractionation. Gel slices containing the whole IP products were then excised and submitted to an in-gel tryptic digestion and mass spectrometry analysis by the Proteomics facility at the Carlos Chagas Institute, as previously described (27). To confirm the specificity of the IP assays, for each polypeptide, the ratio between the intensity generated from the IPs using the extracts expressing the HA-tagged ZC3H41 and the intensity from the control IP using an extract from non-transfected cells was first determined. The base 2 logarithms of the values produced were then calculated for two independent experiments carried out with different cytoplasmic extracts and only the polypeptides whose average ratio were >2 were considered. Identified polypeptides were then ranked and listed in Table 1 with the highest-ranking values according to each category on top.

RNA sequencing

The RNAs ligands were extracted from three independent IPs of wild-type *L. infantum* and HA-tagged ZC3H41 with the RNeasy Mini Kit (QIAGEN) according to the manufacturer's instructions. The RNA samples were quantified with the Qubit™ RNA HS Assay Kit (Thermo Fisher) using the Qubit® 2.0 Fluorometer. A sample of 0.1 to 4 µg of the total RNA was used to construct the cDNA libraries with the TrueSeq Stranded mRNA Library Prep Kit (Illumina), following the manufacturer's instructions. The libraries were validated quantitatively for qPCR use the KAPA Library Quantification Kits and qualitatively by visualization in agarose gel. These libraries were then normalized and the pool prepared to sequencing using the MiSeq® Reagent Kit v3, 150 cycle (Illumina). The data from Illumina were analyzed by bioinformatics tools as previously described (unpublished data - Assis et al., 2019). The results with the upregulated mRNAs that come associated with HA-tagged ZC3H41 were plotted on graphs according to GO terms classification. The minimum of 1.5-fold increase over the negative control (Log2 (I Ratio) >0.6) and FDR of 0.05 were considered.

RESULTS

ZC3H41 is constantly expressed in *Leishmania infantum*

Cell extracts of *Leishmania infantum* expressing HA-tagged ZC3H41 were tested with anti-HA antibody by western blot assays. The expression was confirmed, with ZC3H41 migrating in its approximated size of 60 kDa. These cells extracts were also obtained at several time points to produce a parasite growth curve to analyse changes in the expression profile during the adaptation phases (0-6h), exponential growth (24-48h), stationary (72-96h) and early cell death (120h). The growth curve was tested by western blotting first with anti-HA antibody and then with anti-EIF4AI as a loading control. The ZC3H41 protein was found to be constitutively expressed during all points of the curve in its predicted size of 60 kDa, indicating that post-translational modifications do not occur (Figure 1B).

ZC3H41 associates to PABP homologues but not with EIF4AI

To characterize and confirm the association with possible partners, proteins co-precipitated with the HA-tagged ZC3H41 were identified through western blotting and mass-spectrometry. The IP of negative control, a *L. infantum* cell without HA-tagged proteins, confirm the absence of precipitated proteins, while ZC3H41 was found present in the corresponding immunoprecipitated samples through western blotting using the anti-HA antibody. The IPs were also analyzed with anti-*Lm*PABP1, anti-*Lm*PABP2 and anti-*Tb*EIF4AI. A trace of PABP1 and PABP2 were found co-precipitated with the ZC3H41 IP, but no interaction was observed with EIF4AI (Figure 2). To confirm these results, the IPs were analyzed by mass spectrometry and Table 1 summarizes the results derived from two sets of replicates comparing the LFQ intensity of ZC3H41-HA with the corresponding negative control. These results showed all three PABP homologues co-precipitated with ZC3H41. Interestingly, PABP2 and PABP3 are able to interact directly with each other (6) and with the same mRNAs (unpublished data - Assis et al., 2019)**.** The results also confirmed a lack of association of ZC3H41 with EIF4AI as seen in the western blot assays and also a lack of association with eIF4F-like complexes previously shown to co-precipitate with PABP1 (21).

ZC3H41 also associates to RNA binding proteins

Associations of ZC3H41 with several RBPs or RBP partners are also observed in the mass spectrometry analysis, such as RBP23, LinJ.18.0300 and LinJ.05.0450 (Table 1). These three proteins, plus ZC3H41, are found specifically co-precipitated with PABP1 (21), and the LinJ.18.0300, LinJ.05.0450, ZC3H41and PABP1 also co-precipitate with RBP23, this one shown to directly interact with PABP1 (unpublished data - Assis et al., 2019). Table 1 also presents three more RNA-related proteins co-precipitate exclusively with ZC3H41: (1) LinJ.27.2020, an RNA binding protein with both RRM and zinc finger CCHC-type domains, and zinc ion or nucleic acid binding function; (2) LinJ.07.1020, a splicing factor TSR1-like protein, that in *T. brucei* contain a RRM domain; (3) LinJ.27.0980, a mitochondrial RNA binding protein 1 with an MRP domain and regulation of transcription function. Interestingly, two mitochondrial enzymes also co-precipitated with ZC3H41, a Pyruvate dehydrogenase E1 component alpha subunit and a Dihydrolipoamide acetyltransferase precursor, both belonging to the pyruvate dehydrogenase complex which catalyse the conversion of pyruvate to acetyl-CoA and CO_2, with molecular function of oxidoreductases and transferases, respectively.

mRNA populations associated with ZC3H41

After the IPs, the mRNAs bound to ZC3H41-HA were extracted and used to construct cDNA libraries and next generation sequencing. The negative control used was cytoplasmic extracts from cells not expressing the HA-tagged protein. The mRNAs identified upregulated, which means twofold more abundant than in the negative control, were grouped manually according their molecular function description: binding, catalytic activity, structural constituent of ribosome, transcription regulator activity translation regulator activity, transporter activity. A total of 94 transcripts were found co-precipitated with ZC3H41 of which 4 encode proteins with unknown function (Figure 3). Of the functional mRNAs, 67 (74%) encode ribosomal proteins, a result similar to the one found for the ZC3H41 partners proteins PABP1 and RBP23, followed by 8 (9%) histones mRNAs (Figure 3), which are found more frequently co-precipitated with *Lm*PABP2-PABP3 (unpublished data - Assis et al., 2019). In addition, among the enriched mRNAs, 7 (8%) encode proteins with catalytic activity, 4 (4%) with binding function (DNA binding not included), 1 (1%) with

transcription regulator activity, 1 (1%) with translation regulator activity and 2 (2%) with transporter activity (2%) (Figure 3).

Polyclonal antibody against ZC3H41 is sensitive and specific

The recombinant His-tagged ZC3H41 protein was used to immunize rabbits to obtain polyclonal sera. The purified antibodies were tested with *Escherichia coli* extract expressing recombinant Glutathione S-Transferase (GST)-tagged ZC3H41 and a negative control expressing LinJ.33.0390 to evaluate the specificity. The recombinant GST-tagged ZC3H41 was recognized by anti-ZC3H41 in the expected size of approximately 86 kDa (60 kDa of the protein plus 26 kDa of the GST). The GST-tagged LinJ.33.0390 protein, used as negative control, was not recognized, confirming the specificity of the antibody produced (Supplementary Figure 1). The antibodies were also tested with different concentrations of whole *L. infantum* extracts to evaluate the sensitivity. The anti-ZC3H41 recognized the wild type protein in *Leishmania* extract (data not showed) in the size of 60 kDa. Next step is using these antibodies in localization assays in *L. infantum* cells with and without different inhibitors.

DISCUSSION

ZC3H41 (LinJ.27.1220) (Kramer et al., 2010) share partial similarities with DHH1 and with the eIF4AI and eIF4AIII helicases, mostly due to a helicase conserved C-terminal domain. *L. infantum* promastigotes expressed HA-tagged ZC3H41 constitutively and with no indication of post-translational modifications. New tests with anti-ZC3H41, which has been showed to be efficient and specific, should be performed to analyse whether the endogenous protein presents similar profile. In *T. brucei* silencing the ZC3H41 gene showed that it is essential for trypanosome survival (28). A similar helicase protein, EIF4AI, has a constitutive expression during the parasite life cycle and is essential for protein synthesis and viability (29).

In *T. brucei* ZC3H41 is a protein distributed throughout the cytoplasm but migrates to granules under stress (30). Cytoplasmic spliced leader (SL) RNA is bound by distinct subset of proteins. Among them, ZC3H41 was chosen as the marker for SL RNA containing granules under different conditions (28, 31). To our

knowledge, there is no data about the distribution of ZC3H41 in the *Leishmania* genus, so the antibodies produced here will be used in regular subcelular localization assays and under different inhibitors conditions.

ZC3H41 is associated with PABP1, RBP23, NFT2-like protein LinJ.18.0300 and a hypothetical protein LinJ.05.0450. Evidence shows that RBP23 interact with the 3' UTR of mRNAs and also directly to PABP1, suggesting a mechanism of mRNAs selections to be translated by the complex EIF4E4/EIF4G3 (Figure 4A) (unpublished data - Assis et al., 2019). Interestingly, the six most enriched proteins in the *T*. *brucei* PABP1 pulldown was the proteins eIF4E4, eIF4G3, PABP1, the RNA binding protein RBP23, a hypothetical protein Tb927.7.7460 (LinJ.05.0450 homolog) and the CCCH type zinc finger protein ZC3H41 (9), a group very similar to that of the *L. infantum* homologue.

The set of more enriched mRNAs co-precipitated with ZC3H41 as well as for PABP1 and RBP23 are composed basically of ribosomal protein mRNAs (unpublished data - Assis et al., 2019), which may indicate this mechanism of selection is specific for this mRNAs and is aided by other RNA binding protein, maybe the same partners found in the mass spectrometry analysis of ZC3H41, PABP1 (21) and RBP23 (unpublished data - Assis et al., 2019): NFT2-like protein LinJ.18.0300 and the hypothetical protein LinJ.05.0450. Although, further characterization for both proteins is required.

ZC3H41 is also associated with PABP2 and PABP3, which interact with each other, are involved in translation and have a possible nuclear function (6, 32), suggesting a complex different from that formed by PABP1 (Figure 4B). Our data are in agreement with those found for *L. infantum* Alba1 and Alba3 which were found associated to ZC3H41 and to three PABPs homologs in both developmental stages (33). Protein-protein interaction assays should be performed to further characterize these new complexes.

REFERENCES

1. Jackson,R.J., Hellen,C.U.T. and Pestova,T. V (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat. Rev. Mol. Cell Biol.*, **11**, 113–127.

2. Aitken,C.E. and Lorsch,J.R. (2012) A mechanistic overview of translation initiation in eukaryotes. *Nat. Struct. Mol. Biol.*, **19**, 568–576.

3. Merrick,W.C. (2015) eIF4F: A retrospective. *J. Biol. Chem.*, **290**, 24091–24099.

4. Hinnebusch,A.G. (2014) The Scanning Mechanism of Eukaryotic Translation Initiation. *Annu. Rev. Biochem.*, **83**, 779–812.

5. Haile,S. and Papadopoulou,B. (2007) Developmental regulation of gene expression in trypanosomatid parasitic protozoa. *Curr. Opin. Microbiol.*, **10**, 569–577.

6. da Costa Lima,T.D., Moura,D.M.N., Reis,C.R.S., Vasconcelos,J.R.C., Ellis,L., Carrington,M., Figueiredo,R.C.B.Q. and de Melo Neto,O.P. (2010) Functional characterization of three *Leishmania* poly(A) binding protein homologues with distinct binding properties to RNA and protein partners. *Eukaryot. Cell*, **9**, 1484–1494.

7. Moura,D.M.N., Reis,C.R.S., Xavier,C.C., da Costa Lima,T.D., Lima,R.P., Carrington,M. and De Melo Neto,O.P. (2015) Two related trypanosomatid eIF4G homologues have functional differences compatible with distinct roles during translation initiation. *RNA Biol.*

8. De Melo Neto,O.P., Da Costa Lima,T.D.C., Xavier,C.C., Nascimento,L.M., Romão,T.P., Assis,L. a., Pereira,M.M.C., Reis,C.R.S. and Papadopoulou,B. (2015) The unique *Leishmania* EIF4E4 N-terminus is a target for multiple phosphorylation events and participates in critical interactions required for translation initiation. *RNA Biol.*, **12**, 1209–1221.

9. Zoltner,M., Krienitz,N., Field,M.C. and Kramer,S. (2018) Comparative proteomics of the two *T. brucei* PABPs suggests that PABP2 controls bulk mRNA. *PLoS Negl. Trop. Dis.*, **12**, 1–18.

10. Yoffe,Y., Léger,M., Zinoviev,A., Zuberek,J., Darzynkiewicz,E., Wagner,G. and Shapira,M. (2009) Evolutionary changes in the *Leishmania* eIF4F complex involve variations in the eIF4E-eIF4G interactions. *Nucleic Acids Res.*, **37**, 3243–3253.

11. Yoffe,Y., Zuberek,J., Lerer,A., Lewdorowicz,M., Stepinski,J., Altmann,M., Darzynkiewicz,E. and Shapira,M. (2006) Binding specificities and potential roles of isoforms of eukaryotic initiation factor 4E in *Leishmania*. *Eukaryot. Cell*, **5**, 1969–1979.

12. Zinoviev,A., Léger,M., Wagner,G. and Shapira,M. (2011) A novel 4E-interacting protein in *Leishmania* is involved in stage-specific translation pathways. *Nucleic Acids Res.*, **39**, 8404–8415.
13. Takeuchi,O. and Yamashita,A. (2017) Translational control of mRNAs by 3'-Untranslated region binding proteins. *BMB Rep.*, **50**, 194–200.
14. Kramer,S. and Carrington,M. (2011) Trans-acting proteins regulating mRNA maturation, stability and translation in trypanosomatids. *Trends Parasitol.*, **27**, 23–30.
15. De Gaudenzi,J.G., Carmona,S.J., Agüero,F. and Frasch,A.C. (2013) Genome-wide analysis of 3'-untranslated regions supports the existence of post-transcriptional regulons controlling gene expression in trypanosomes. *PeerJ*, **1**, 1–28.
16. Müller-McNicoll,M. and Neugebauer,K.M. (2013) How cells get the message: dynamic assembly and function of mRNA-protein complexes. TL - 14. *Nat. Rev. Genet.*, **14 VN-r**, 275–287.
17. Oliveira,C., Faoro,H., Alves,L.R. and Goldenberg,S. (2017) RNA-binding proteins and their role in the regulation of gene expression in *Trypanosoma cruzi* and *Saccharomyces cerevisiae*. *Genet. Mol. Biol.*, **40**, 22–30.
18. Wang,D., Guo,Y., Wu,C., Yang,G., Li,Y. and Zheng,C. (2008) Genome-wide analysis of CCCH zinc finger family in *Arabidopsis* and rice. *BMC Genomics*, **9**, 1–20.
19. Berg,J.M. and Shi,Y. (2012) The Galvanization of Biology : Appreciation for the Roles of Growing Zinc. *Adv. Sci.*, **271**, 1081–1085.
20. Schmitt,E., Naveau,M. and Mechulam,Y. (2010) Eukaryotic and archaeal translation initiation factor 2: A heterotrimeric tRNA carrier. *FEBS Lett.*, **584**, 405–412.
21. de Melo Neto,O.P., da Costa Lima,T.D.C., Merlo,K.C., Romão,T.P., Rocha,P.O., Assis,L.A., Nascimento,L.M., Xavier,C.C., Rezende,A.M., Reis,C.R.S., *et al.* (2018) Phosphorylation and interactions associated with the control of the *Leishmania* Poly-A Binding Protein 1 (PABP1) function during translation initiation. *RNA Biol.*, **15**, 739–755.
22. Kramer,S., Carrington,M., Article,I., Url,A., Central,P. and Central,B. (2010) Genome-wide *in silico* screen for CCCH-type zinc finger proteins of. *BMC*

Genomics, **11**, 283.

23. Singh,A., Minia,I., Droll,D., Fadda,A., Clayton,C. and Erben,E. (2014) Trypanosome MKT1 and the RNA-binding protein ZC3H11: Interactions and potential roles in post-transcriptional regulatory networks. *Nucleic Acids Res.*, **42**, 4652–4668.
24. Padmanabhan,P.K., Dumas,C., Samant,M., Rochette,A., Simard,M.J. and Papadopoulou,B. (2012) Novel features of a PIWI-like protein homolog in the parasitic protozoan *Leishmania*. *PLoS One*, **7**, e52612.
25. De Melo Neto,O.P., Standart,N. and De Sa,C.M. (1995) Autoregulation of poly(A)-binding protein synthesis *in vitro*. *Nucleic Acids Res.*, **23**, 2198–2205.
26. Mureev,S., Kovtun,O., Nguyen,U.T.T. and Alexandrov,K. (2009) Species-independent translational leaders facilitate cell-free expression. *Nat. Biotechnol.*, **27**, 747–752.
27. Rezende,A.M., Assis,L.A., Nunes,E.C., da Costa Lima,T.D., Marchini,F.K., Freire,E.R., Reis,C.R. and de Melo Neto,O.P. (2014) The translation initiation complex eIF3 in trypanosomatids and other pathogenic excavates--identification of conserved and divergent features based on orthologue analysis. *BMC.Genomics*, **15**, 1175.
28. Eliaz,D., Kannan,S., Shaked,H., Arvatz,G., Tkacz,I.D., Binder,L., Waldman Ben-Asher,H., Okalang,U., Chikne,V., Cohen-Chalamish,S., *et al.* (2017) Exosome secretion affects social motility in *Trypanosoma brucei*.
29. Dhalia,R., Marinsek,N., Reis,C.R.S., Katz,R., Muniz,J.R.C., Standart,N., Carrington,M. and de Melo Neto,O.P. (2006) The two eIF4A helicases in *Trypanosoma brucei* are functionally distinct. *Nucleic Acids Res.*, **34**, 2495–2507.
30. Fritz,M., Vanselow,J., Sauer,N., Lamer,S., Goos,C., Siegel,T.N., Subota,I., Schlosser,A., Carrington,M. and Kramer,S. (2015) Novel insights into RNP granules by employing the trypanosome's microtubule skeleton as a molecular sieve. *Nucleic Acids Res.*, **43**, 8013–8032.
31. Dean,S., Sunter,J.D. and Wheeler,R.J. (2017) TrypTag.org: A Trypanosome Genome-wide Protein Localisation Resource. *Trends Parasitol.*, **33**, 80–82.
32. Kramer,S., Bannerman-Chukualim,B., Ellis,L., Boulden,E.A., Kelly,S., Field,M.C. and Carrington,M. (2013) Differential Localization of the Two *T. brucei* Poly(A)

Binding Proteins to the Nucleus and RNP Granules Suggests Binding to Distinct mRNA Pools. *PLoS One*, **8**.

33. Dupé,A., Dumas,C. and Papadopoulou,B. (2015) Differential subcellular localization of *Leishmania* alba-domain proteins throughout the parasite development. *PLoS One*, **10**, 1–29.

Table 1. Proteins co-purifying with HA-tagged ZC3H41.

Identificação	Gene ID	Log_2 (I^a $Ratio^b$)		Media>1,9
		1st Exp	2nd Exp	
Translation				
PABP2	LinJ.35.4200	1,30	3,49	2,39
PABP3	LinJ.25.0080	1,09	3,56	2,32
PABP1	LinJ.35.5360	1,11	2,78	1,94
RNA related				
ZC3H41	LinJ.27.1220	4,46	6,24	5,35
RNA binding protein – RBP23	LinJ.17.0610	0,77	4,50	2,63
RNA binding protein – RBP21	LinJ.27.0980	1,04	3,63	2,33
Splicing factor TSR1	LinJ.07.1020	0,86	3,72	2,29
RNA binding protein	LinJ.27.2020	0,57	3,87	2,22
Catalytic activity				
Pyruvate dehydrogenase E1 component alpha subunit	LinJ.18.1360	2,24	4,73	3,48
Dihydrolipoamide acetyltransferase precursor	LinJ.36.2790	2,40	2,82	2,61
Uncharacterized protein				
Uncharacterized protein	LinJ.05.0450	3,68	6,05	4,86
Uncharacterized protein containing NTF2 and RRM-like domain	LinJ.18.0300	1,48	3,38	2,43

*11 polypeptides are shown which specifically co-immunoprecipitated with the HA-tagged ZC3H41 in two independent immunoprecipitation (IP) experiments, with a minimum of 8-fold increase over the negative control (Log_2 (I Ratio) >2).

[a]I= Intensity

[b]R = Ratio between the intensity observed for the IP with extracts derived from the cells expressing HA tagged ZC3H41 divided by the intensity detected for IP with the extracts from wild type cells.

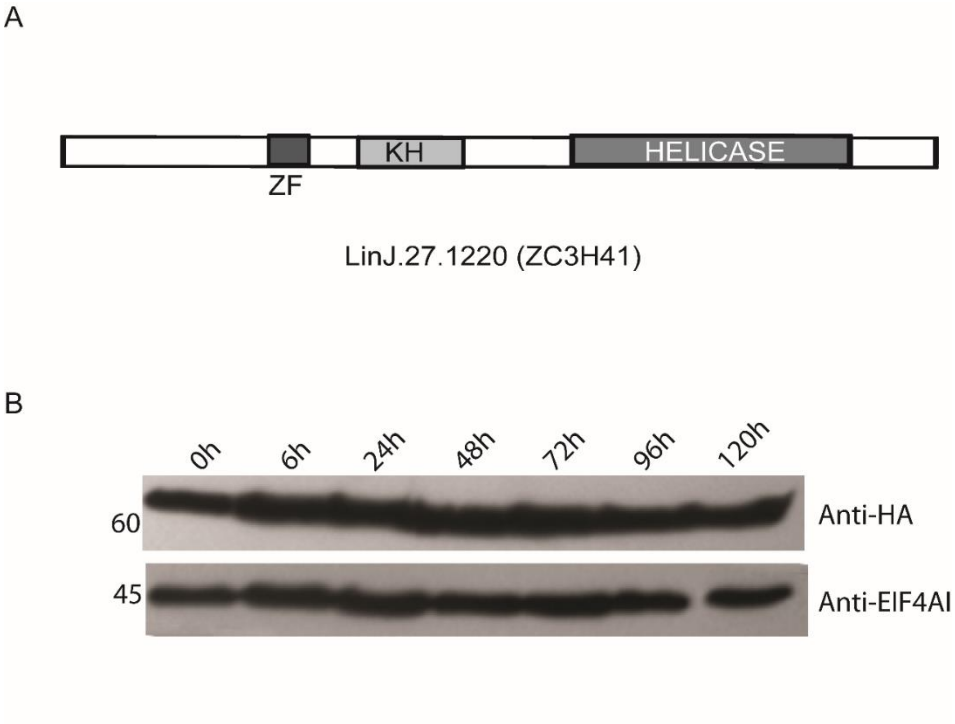


Figure 1. ZC3H41 protein expression profile during the life cycle of *L. infantum*. A) Schematic representation of *L. infantum* ZC3H41 (LinJ.27.1220) with a CCCH-type Zinc Finger domain (ZF), a K homolog (KH) domain and a helicase C-terminal conserved domain (HELICASE). **B)** Anti-HA monoclonal antibody was used to detect the HA-tagged ZC3H41 and the anti-TbEIF4AI polyclonal antibody as load control, both diluted 1:3000. The second antibody used was anti-mouse IgG diluted 1: 3000 for monoclonal and anti-rabbit IgG diluted 1: 10000 for polyclonal.

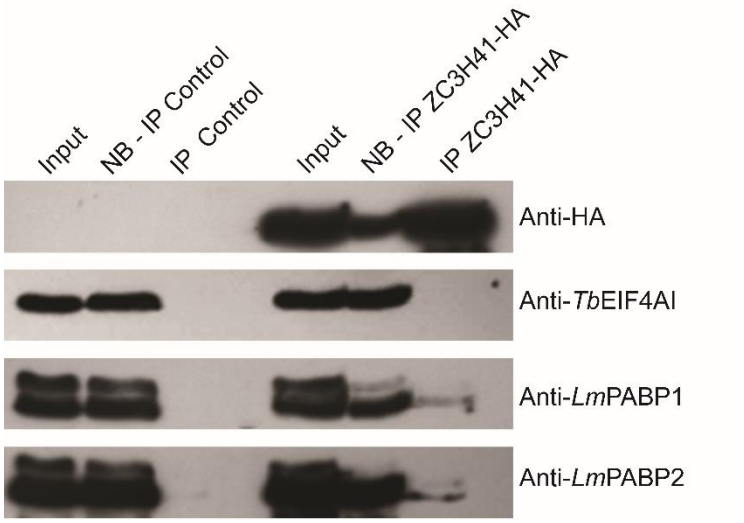


Figure 2. Western blot assay of the HA-tagged ZC3H41 protein immunoprecipitations. **Line 1** - Anti-HA monoclonal antibody (1:3000) detected the 60 kDa size protein in cytoplasmic extract (input), non-bound (NB) and immunoprecipitated (IP). No recognition to IP control as expected. The second antibody was anti-mouse IgG diluted 1: 3000. **Line 2** - Polyclonal anti-TbEIF4AI (1:3000). **Line 3** - Polyclonal anti-LmPABP1 (1:2000). **Line 4** - Polyclonal anti-LmPABP2 (1:2000). The second antibody used for polyclonal antibodies was anti-rabbit IgG diluted 1: 10000. The polyclonal antibodies detected endogenous protein in the input and NB fractions, and in IP fraction only when is co-precipitated with HA-tagged protein. It was detected PABP1 and PABP2 but not EIF4AI co-precipitated with HA-tagged ZC3H41.

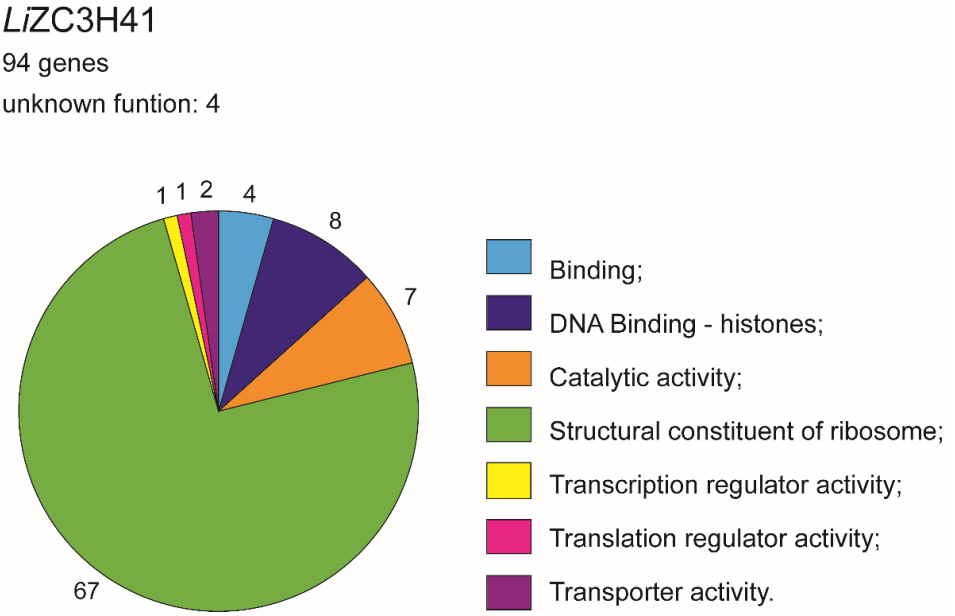


Figure 3. Analysis of mRNAs populations associated with LiZC3H41. The upregulated genes (at least twofold enrichment in comparison with the negative control) were grouped manually according to functional categories (Binding; DNA binding – histones; Catalytic activity; Structural constituent of ribosome; Transcription regulator activity, Translation regulator activity, Transporter activity). Genes encoding proteins of unknown function are not shown.

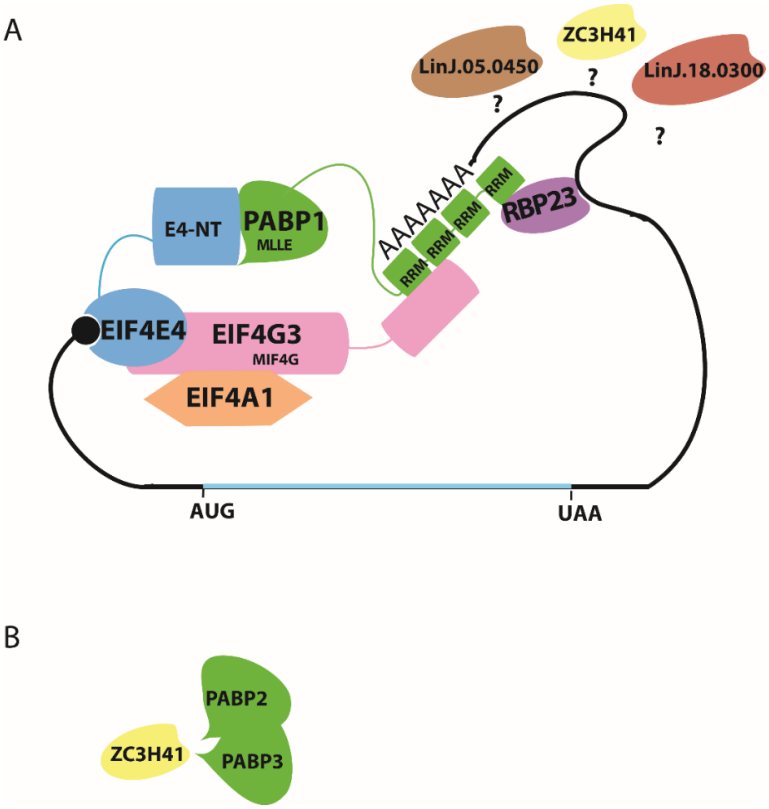
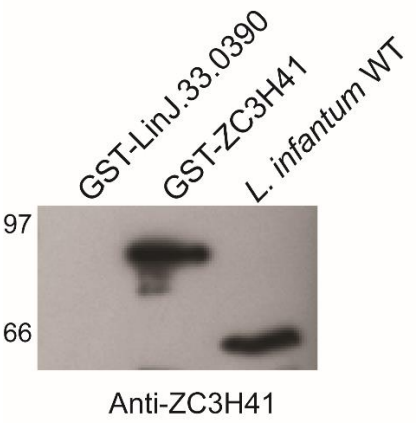


Figure 4. Two different complexes for ZC3H41. A) ZC3H41 interacts with PABP1 and other proteins as RBP23, LinJ.18.0300 and LinJ.05.0450 (RBP23 complex). PABP1 bound to poly(A) tail interacts directly with RBP23 complex; RBP23 complex interacts with 3'UTR motif on the ribosomal protein mRNA; the RBP23 complex recruits EIF4E4/EIF4G3/EIF4AI via PABP1-EIF4G3 interaction; **B)** An independent complex with PABP2 and PABP3 and unclear function. Adapted from: unpublished data - Assis et al., 2019.



Supplementary Figure 1. Western blot assays with the polyclonal anti-*Li*ZC3H41. The purified antibodies (1:1500) were used against wild-type extract of *L. infantum* (~60 kDa) to verify sensitivity and against purified protein GST-tagged ZC3H41 (~86 kDa) and GST-tagged LinJ.33.0390, which was not detected indicating the specificity of the antibody produced.

5 ARTIGO 3

5.1 Avaliação do RNA de interferência para RBP23 e proteínas parceiras em *Trypanosoma brucei.*

Artigo anexo na página seguinte.

A doutoranda e primeira autora desse artigo em redação e em processo de finalização foi responsável pela escrita e por toda a experimentação sob a orientação do Dr. Mark Carrington durante o doutorado sanduíche no Departamento de Bioquímica da Universidade de Cambridge, Reino Unido. Resumidamente, as análises após RNAi do crescimento de células procíclicas expressando a RBP23 e proteínas parceiras marcada com eYFP não mostraram diferença significativa em relação a curva controle. Através de citometria de fluxo e microscopia foi confirmada que uma linhagem teve a proteína depletada das células, duas tiveram a quantidade reduzida, porém a RBP23 permaneceu intacta. Um segundo método foi utilizado para a RBP23 em células procíclicas e sanguínea e novamente nenhuma diferença significante no crescimento celular foi observada. As ferramentas obtidas devem em futuro próximo permitir uma finalização no que concerne a análise de parceiros de cada uma das proteínas investigadas e avaliando a existência de um complexo maior baseado na RBP23.

Avaliação por RNA de interferência da RBP23 e proteínas parceiras em *Trypanosoma brucei*

Ludmila A. Assis[1,2], Osvaldo P. de Melo Neto[1] e Mark Carrington[3]

[1] Departamento de Microbiologia, Instituto Aggeu Magalhães – FIOCRUZ, Recife, PE, Brasil

[2] Universidade Federal de Pernambuco, Recife, PE, Brasil

[3] Departamento de Bioquímica, Universidade de Cambridge, Cambridge, Reino Unido

Palavras-chave: Tripanosomatidae. RBP23. PABP1.

RESUMO

Em tripanossomatídeos, a RBP23 é uma proteína de ligação a RNA que interage quase que exclusivamente com mRNAs codificantes de proteínas ribossomais. Dessa forma a RBP23 parece auxiliar na regulação da tradução desses mRNAs através da sua interação direta com a PABP1, homólogo da proteína de ligação ao poli-A, e com o complexo de iniciação da tradução eIF4F formado pelas subunidades EIF4E4/EIF4G3/EIF4AI. Quatro proteínas ainda não caracterizadas foram encontradas associadas de forma específica a RBP23 de *Leishmania infantum* e podem auxiliar sua função. Nesse trabalho ortólogos desse conjunto de proteínas não caracterizadas, bem como da própria RBP23, foram selecionadas para uma melhor caracterização, usando como modelo o *Trypanosoma brucei*. Para isso, o gene endógeno que codifica cada uma dessas proteínas foi marcado com fluorescência. As linhagens celulares resultantes foram então transfectadas com plasmídeos que permitem a depleção por RNAi das respectivas proteínas fluorescentes. Não houve diferenças significativas no crescimento celular das linhagens procíclicas com adição de doxiciclina, que é um indutor de RNAi. Diferentes métodos foram utilizados para confirmar a depleção das proteínas como citometria de fluxo e microscopia. Uma linhagem teve a proteína depletada das células, duas tiveram a sua quantidade reduzida, porém a RBP23 permaneceu

intacta. Por isso, um segundo método foi utilizado para depletar a RBP23 e novamente nenhuma diferença significante no crescimento celular foi observada para células procíclicas e sanguíneas. Mais estudos deverão ser realizados para confirmação da nova técnica, e dessa forma utilizar essas linhagens para estudos mais avançados de caracterização funcional.

INTRODUÇÃO

O *Trypanosoma brucei*, organismo responsável pela Tripanossomíase Humana Africana, ou Doença do Sono, possui duas principais formas: sanguínea no hospedeiro mamífero e procíclica no inseto vetor. Similar a outros tripanossomatídeos, a regulação da expressão dos seus genes ocorre pós-transcricionalmente, devido à falta de promotores para a RNA polimerase II (1). Proteínas de ligação ao RNA (RBPs) influenciam a estabilidade e a tradução de mRNAs, atuando como reguladores, ativadores ou repressores da expressão gênica de tripanossomatídeos (2–4). Em *T. brucei* foi demonstrado que a RBP23, uma destas proteínas, é capaz de aumentar a expressão de um mRNA repórter, considerada assim um ativador da tradução. O mesmo é observado para a PABP1, homólogo da proteína de ligação ao poli-A, e para as subunidades do complexo do tipo eIF4F (EIF4E4, EIF4G3 e EIF4AI), parceiro desta proteína. (5). Esse complexo, junto com a PABP1 é responsável pela tradução de parte significativa dos mRNAs de tripanossomatídeos (6, 7).

Em *Leishmania infantum* um novo modelo foi sugerido de regulação da tradução, no qual a RBP23, associada em sua maioria a mRNAs codificantes de proteínas ribossomais, participa da seleção dos mesmos para a tradução pelo complexo EIF4E4/EIF4G3/EIF4AI, via interação direta com a PABP1. Além da interação com este complexo, a RBP23 também interage de forma específica com proteínas ainda não caracterizadas (LinJ.21.1510, LinJ.040140, LinJ.19.0020 e LinJ.05.0450) que podem auxiliar em sua função (Assis *et al*., – manuscrito em preparação). Interessantemente, a quantidade de RBP23 diminui durante a diferenciação e nas formas amastigotas (8), o que pode significar uma função estágio dependente.

Espécies modelo de *Leishmania* não respondem a RNAs dupla fita (dsRNA), e, portanto, não possuem uma via funcional de RNA de interferência (RNAi) que facilitem uma melhor caracterização funcional de proteínas específicas (9). Nesse trabalho, como alternativa, os ortólogos de *T. brucei* da RBP23 e das suas quatro proteínas parceiras (Tb927.10.810, Tb927.9.7080, Tb927.10.14700 e Tb927.7.7460) foram selecionados para uma melhor caracterização, pois nesse organismo é possível utilizar o RNAi, além de outras técnicas de validação funcional. Assim, as formas endógenas cada uma das proteínas selecionadas, e da própria RBP23, foram primeiramente modificadas em células procíclicas através da marcação com a proteína fluorescente eYFP. As linhagens celulares resultantes foram então utilizadas em experimentos de indução de dsRNA e RNAi, sendo seu efeito avaliado por curva de crescimento, citometria de fluxo e microscopia. Por fim, uma segunda técnica mais estável (10) que expressa o RNAi em forma de grampo, foi utilizada apenas para a RBP23 em células procíclicas e sanguíneas. Foram geradas linhagens que permitiram analisar a localização e a influência dessas proteínas no crescimento celular, e que depois de melhor padronizadas podem servir para estudos futuros experimentos de interação com o complexo EIF4E4/EIF4G3/EIF4AI e a PABP1 além de seus mRNAs alvos.

MATERIAIS E MÉTODOS

Construção dos plasmídeos e manipulação de DNA

Para marcar os genes endógenos com eYFP foram desenhados *primers* para amplificar plasmídeos contendo tanto o gene eYFP como genes de resistência a antibióticos como marcas de seleção. Os genes *Tb927.10.810*, *Tb927.9.7080* e *Tb927.10.14700* foram marcados no N-terminal enquanto os genes *RBP23* e *Tb927.7.7460* no C-terminal. Para marcar o N-terminal foi utilizado o plasmídeo 2827 como molde para o iniciador 5’, compreendendo os 80 nucleotídeos imediatamente antes da região codificadora (CDS) do gene alvo mais 20 nucleotídeos do início da CDS do gene de resistência a blasticidina. O iniciador 3' corresponde aos 20 nucleotídeos finais da CDS do gene eYFP e 80 nucleotídeos iniciais da CDS do gene alvo. Para o C-terminal foi utilizado o plasmídeo 2710 como molde para o iniciador 5', consistindo em 80 nucleotídeos do fim da CDS do gene alvo sem o

códon de parada mais 20 nucleotídeos iniciais da CDS do gene eYFP. O iniciador 3' com 20 os nucleotídeos finais da CDS do gene de resistência a G418 mais os 80 nucleotídeos seguintes a CDS do gene alvo. Os iniciadores estão representados na Tabela 1). Todas as PCR foram realizadas segundo recomendações do fabricante do Expand™ High Fidelity PCR System. Os fragmentos obtidos foram transfectados em células de *T. brucei* que já continham o plasmídeo pSMOX que expressa a T7 polimerase (T7pol) e o repressor tetraciclina (TetR).

Para obtenção dos plasmídeos que permitem a produção do RNA de interferência, o DNA genômico de *T. brucei*, cepa Lister 427 KG foi isolado utilizando DNeasy Blood & Tissue Kits (Qiagen), seguindo as recomendações do fabricante. Os genes *RBP23, Tb927.10.810; Tb927.9.7080; Tb927.10.14700* e *Tb927.7.7460* foram amplificados utilizando iniciadores flanqueados pelas enzimas de restrição HindIII e BamHI (ver Tabela 1). Todos os fragmentos de PCR foram primeiro clonados no vector pGEM-T Easy (Promega) para sequenciamento e subclonados nos mesmos sítios de clonagem, HindIII-BamHI, do vector de expressão p2T7-177, que possui dois promotores T7 opostos, permitindo a formação de um dsRNA. Os plasmídeos foram transfectados nas respectivas células marcadas com eYFP.

Para obtenção do RNAi em forma de grampo o gene da RBP23 HindIII-BamHI foi clonado duas vezes no plasmídeo 3666 (10). Primeiro foi feita a clonagem no sítio BamHI-AarI, e uma vez o plasmídeo recuperado esse foi cultivado em células de linhagem *dam* negativa que removem a metilação do segundo sítio de clonagem. Em seguida outro fragmento da RBP23 foi clonado no sítio HindIII-BclII. A construção foi transfectada em células procíclicas e sanguíneas contendo o plasmídeo pSMOX. Todos os produtos de plasmídeo foram linearizados com NotI-HF antes das transfecções.

Crescimento dos parasitas e análise de expressão

Células procíclicas de *Trypanosoma brucei* da linhagem Lister 427 KG pSMOX foram cultivadas em meio SDM-79 suplementado com 10% de Soro Fetal Bovino (SFB) inativado e 3% de hemina a 27ºC e diluídas diariamente para $1x10^6$ células/ml. As células sanguíneas da linhagem Lister 427 KG foram cultivadas a 37°C em HMI-9 suplementado com 10% de SFB e diluídas diariamente para $1x10^5$ células/ml. Os procedimentos de transfecção utilizados para produtos de PCR e/ou

plasmídeos linearizados foram realizados por eletroporação. Resumidamente, 10 µg de DNA em tampão 3xR (200 mM de Na_2HPO_4, 70 mM de NaH_2PO_4, 15 mM de KCl, 150mM de HEPES pH 7,3 e 1,5mM de $CaCl_2$) foram utilizados para ressuspender 3-6x10^7 células centrifugadas. Estas foram em seguida transferidas para cuvetas de 2 mm que foram submetidos a Gene Pulser II (Bio-Rad) com pulso de 1,5 V, 25 µF. Foram feitas três diluições em série e após 6 horas as células foram selecionadas com antibióticos (ver Tabela 2). As células foram distribuídas em placas de 24 poços e os clones foram selecionados em torno do décimo dia após a transfecção.

Imagens de microscopia

Para visualizar a expressão eYFP em células transfectadas, células procíclicas foram cultivadas para 5x10^6 a 10^7 células/ml, com 5 µl destas colocados em uma lâmina SuperFrost® plus com uma lamínula de 25x50mm, com visualização em microscópio de fluorescência (Zeiss Ax10, objetiva 100x com óleo). As imagens foram capturadas usando o CCD AxioCam e analisadas usando o software AxioVision V4.8.2. O tempo de exposição foi fixado de 4000 ms para YFP e 141 ms para DIC.

Imunodetecção

Extratos celulares foram gerados centrifugando-se culturas na fase exponencial e ressuspendendo-se o sedimento diretamente em tampão de amostra para SDS-PAGE (10% de B-mercaptoetanol, 6% de SDS, 240 mM de Tris-HCl pH 6.8, 30% glicerol e 0.006% de azul de bromofenol), para o equivalente a 2x10^8 células/ml. Cerca de 2x10^6 de cada amostra foram fracionadas em 17.5% SDS-PAGE e transferidos para Membrana Immobilon-FL (Millipore), seguindo o protocolo padrão de laboratório (11). Em seguida, a primeira incubação foi realizada com anticorpo policlonal anti-GFP (Molecular Probes) para detectar eYFP em diluição de 1: 1000 e a segunda incubação foi realizada com anti-IgG de coelho conjugado com peroxidase diluído 1: 5000. A detecção foi realizada por quimioluminescência usando filme Fuji Medical X-Ray. Os filmes foram digitalizados e os sinal foi quantificado usando ImageJ 1.48v. O marcador pré-corado utilizado foi o Precision Plus Protein™ (Bio-Rad).

RNAi

O plasmídeo de RNAi p2T7-177 foi selecionado após transfecção utilizando fleomicina (2,5 mg/ml) e o p3666 utilizando blasticidina (5 mg/ml). As células procíclicas transfectadas com p2T7-177 ou com p3666 foram diluídas para 10^6 células/ml e a forma sanguínea que foi transfectada apenas com p3666 a 10^5 células/ml. Em seguida, adicionou-se 1 mg/ml do antibiótico doxiciclina a culturas no meio da fase exponencial para indução do RNAi. Essas células foram monitoradas entre 3 e 5 dias e diluídas quando necessário. A curva de controle para cada linhagem celular foi feita com sua respectiva célula sem doxiciclina. A contagem de cada dia foi plotada em gráficos, nos quais 10^6 foram considerados 1 para procíclica e 10^5 para sanguínea. O eixo Y foi apresentado em escala logarítmica e o eixo X com intervalos de 24 horas.

Citometria de fluxo

As células fluorescentes e uma linhagem celular não fluorescente foram analisadas por citometria de fluxo após indução de RNAi durante 5 dias. As células procíclicas foram cultivadas até o meio da fase exponencial quando foi adicionado 1 mg/ml de doxiciclina. Após 24 h, as células foram contadas e diluídas para 10^6 células/ml, sendo novamente adicionada doxiciclina, com 1 ml de cada cultura de células analisada no citômetro de fluxo. Os dados foram adquiridos usando o FACScan (Becton Dickinson) e analisados usando o software Cell Quest V3.3 e FlowJo v10.

RESULTADOS

Silenciamento da RBP23 e proteínas parceiras

Dados anteriores de espectrometria de massa de RBP23 em *Leishmania infantum* identificaram quatro proteínas parceiras não caracterizadas (Assis *et al*., – manuscrito em preparação). Os ortólogos de *T. brucei* destes parceiros (Tb927.10.810, Tb927.9.7080, Tb927.10.14700, Tb927.7.7460) bem como o ortólogos da RBP23 foram inicialmente marcados com a proteína fluorescente eYFP (*enhanced Yellow Fluorescent Protein*). Iniciadores específicos para cada gene foram desenhados e os fragmentos obtidos por PCR, descritos na Figura 1A-B,

foram transfectados em células procíclicas de *T. brucei* contendo o plasmídeo pSMOX, que expressam a T7pol e o TetR. A confirmação da marcação foi observada por microscopia e ensaio de *western blot* (Figura 1C-D). Uma vez confirmada, as linhagens celulares fluorescentes foram utilizadas em novas transfecções, sendo que para cada proteína marcada foi feita a transfecção com o plasmídeo p2T7-177 contendo o seu respectivo gene. Este plasmídeo permite a expressão de RNA dupla fita (dsRNA) do gene clonado na presença do indutor doxiciclina. Quatro das cinco linhagens celulares que expressaram fusões com eYFP foram transfectadas com a respectiva construção com sucesso, com base na seleção positiva de antibiótico.

Crescimento celular após RNAi da RBP23 e proteínas parceiras

Três clones foram recuperados de cada transfecção com o plasmídeo p2T7-177 que foram depois tratados com doxiciclina, o indutor de RNAi. As células foram contadas e analisadas morfologicamente a cada 24 horas após a adição de doxiciclina por um total de cinco dias. Os dados foram plotados em gráficos (Figura 2) em que o eixo Y representa o número de células, o número 1 corresponde ao número inicial de 10^6 células/ml e o eixo X, o número de dias em que a célula foi monitorada. Paralelamente, células cultivadas e monitoradas igualmente, porém sem adição de doxiciclina, foram utilizadas como grupo não tratado. Tanto as células não tratadas como as tratadas dos três clones tiveram uma taxa de crescimento muito semelhante, sem diferença significativa entre elas. Isso pode significar que essas proteínas não são realmente essenciais para o crescimento celular ou que a metodologia escolhida não foi eficiente. Para confirmar uma das duas possibilidades, foi realizada a citometria de fluxo de um clone de cada linhagem celular.

Análise de fluorescência após RNAi da RBP23 e de proteínas parceiras

A citometria de fluxo é uma técnica capaz de detectar e medir a quantidade de fluorescência celular. Espera-se que, após a indução do RNAi, a quantidade de proteína fluorescente diminua. Na Figura 3, a coluna da esquerda mostrou duas células não tratadas: o primeiro pico (púrpura) representa células não fluorescentes enquanto o segundo pico (vermelho) células fluorescentes. Espera-se que, após a

indução do RNAi, o pico se mova da direita para a esquerda, indicando uma diminuição na quantidade de proteína fluorescente.

A proteína Tb927.10.14700 foi a única completamente depletada das células após 48 horas. No entanto, uma vez que nenhuma diferença significativa foi observada no crescimento celular, a proteína não é essencial para a célula. Já a proteína Tb927.9.79780 teve comportamento semelhante; no entanto, não foi completamente esgotada, mostrando que o método para RNAi não foi totalmente eficiente. A proteína Tb927.10.810 tem a menor fluorescência, como visto anteriormente na microscopia (Figura 1C), mas ainda é possível observar que houve uma diminuição quase total na quantidade da proteína. Embora a curva de crescimento não tenha mostrado diferença significativa entre as células tratadas e não tratadas dos três clones, quando foi realizada a curva para o experimento de citometria de fluxo uma diferença significativa após 72 horas foi observada entre as células tratadas e não tratadas (dados não mostrados). No caso da RBP23 não se observou nenhuma alteração na quantidade de proteína ao longo dos dias após a adição de doxiciclina. Nesse caso, como mostrado na Figura 3, todos os picos se sobrepunham, indicando que não houve qualquer redução na quantidade desta proteína.

Como alternativa para se confirmar que o método utilizado para obtenção do RNAi não foi eficiente, foi realizada a microscopia dos três clones de cada linhagem celular. Os três clones de cada linhagem foram visualizados por microscopia com uma exposição fixa a cada 24 horas por até três dias após a adição de doxiciclina. A análise indica que apenas a proteína Tb927.10.14700 foi quase completamente depletada (Figura 4), como visto no resultado de citometria de fluxo. As outras proteínas permaneceram fluorescentes até o último dia e, portanto, uma nova estratégia para a indução de RNAi foi realizada, primeiramente apenas para a RBP23.

Crescimento celular de RBP23 após RNAi alternativo

A indução do RNAi pela expressão de um dsRNA em forma de grampo é mais estável do que experimentos de RNAi utilizando um dsRNA simples. Para isso, contudo, é necessário clonar o mesmo gene duas vezes no plasmídeo p3666 (10). Primeiro, o fragmento RBP23 foi inserido no primeiro local de clonagem

(BamHI/BspMI) e, uma vez confirmado, o plasmídeo com o gene foi cultivado em células dam que remove a metilação do segundo local de clonagem (HindIII/BcII), permitindo a clonagem do segundo fragmento. A construção gerada foi então transfectada em células procíclicas e sanguíneas contendo já transfectadas com o vetor pSMOX, que é necessário para as experiências de RNAi, como mencionado anteriormente. Quatro clones de cada foram recuperados para indução de RNAi, e para o grupo tratado, as células foram contadas durante três dias a cada 24 horas após a adição de doxiciclina. Células sem adição de doxiciclina foram usadas como grupo não tratado. Nenhuma diferença significativa foi observada entre as células tratadas e não tratadas, tanto no crescimento procíclico quanto no sanguíneo (Figura 5). Esses resultados indicam que a proteína parece não ser essencial para a célula, mas ainda são necessários experimentos para confirmar que esse novo método foi eficiente.

DISCUSSÃO

Em *L. infantum*, os ortólogos das proteínas selecionadas para esse estudo interagem com a RBP23 que por sua vez interage majoritariamente com mRNAs de proteínas ribossomais (Assis *et al*., – manuscrito em preparação). Este trabalho buscou entender melhor a função dessas proteínas hipotéticas, avaliando ainda se as mesmas e/ou complexos proteicos associados, bem como os mRNAs de proteínas ribossomais alvos, permaneceriam intactos na ausência de uma ou mais das proteínas selecionadas.

Em estudo prévio em *T. brucei* confirmaram a distribuição citoplasmática da RBP23 pois, mesmo quando submetida a estresse nutricional, permanece no citoplasma e não migra para os grânulos de estresse, algo que também ocorre com outras proteínas associadas a PABP1 (12). Curiosamente os estudos de RNAi para a RBP23 de *T. brucei* são controversos: nos primeiros estudos o RNAi afetou negativamente células sanguíneas (13), porém em estudos mais recentes, as células sanguíneas e durante a diferenciação apresentaram um ganho de *fitness* (14) indicando possivelmente uma alta adaptabilidade em condições adversas. Aqui o primeiro método de RNAi utilizado não foi eficiente, e o segundo demonstrou que

as células sanguíneas e procíclicas não foram afetadas. Porém mais estudos serão necessários para confirmar se a proteína foi realmente depletada destas últimas.

Em *T. brucei*, e assim como ocorre com a RBP23, a proteína Tb927.7.7460 também co-precipita com a PABP1 (12), assim como seus ortólogos em *L. infantum* (Assis *et al*., – manuscrito em preparação). Aqui não foi possível avançar nos ensaios de indução de mRNA desta proteína devido a falhas nas tentativas de transfectar o plasmídeo que expressa dsRNA. Contudo a marcação da proteína eYFP mostrou a proteína distribuída no citoplasma, uma informação ainda não disponível no banco de dados de sublocalização de proteínas de células de *T. brucei,* o *TrypTag*. Em resultados prévios o RNAi dessa proteína afetou negativamente ambas as formas de *T. brucei*: sanguínea e procíclica (14). O ortólogo em *L. infantum* é estruturalmente similar a proteína Skp1 (*S-phase kinase associated protein 1*) uma proteína componente do complexo ubiquitina ligase (15). Esse complexo tem como alvo substratos proteicos para ubiquitinilação e consequentemente degradação pelo proteossoma (16). Outra proteína cuja localização ainda não havia sido publicada, a Tb927.9.7080, foi encontrada nesse estudo distribuída por todo citoplasma. Essa proteína parece ter grande importância na célula uma vez que foi descrito que seu RNAi afeta a forma sanguínea, a procíclica e a diferenciação (14). Existe indícios de que essa proteína seja fosforilada em células procíclicas (17).

Com exceção da proteína Tb927.10.14700, não foi possível deletar completamente as demais proteínas estudadas de suas respectivas células, e com isso não foi possível avançar nos estudos originalmente propostos. A falta da Tb927.10.14700, contudo não afetou o crescimento da célula procíclicas como também observado em outro estudo (14), no qual as células permanecem inalteradas tanto nas formas procíclicas quanto na sanguínea. Essa proteína pode ser encontrada distribuída pelo citoplasma, nucleoplasma e nucléolo (18). Já a Tb927.10.810 é uma proteína citoplasmática, que também foi encontrada na região flagelar do citoplasma e ainda no lúmen nuclear (18), porém aparentemente o RNAi dessa proteína não afeta nenhuma forma do *T. brucei* (14). Essa proteína é classificada como PPR (*Pentatricopetide Repeat protein*) por possuir três motivos desse tipo (19). Em *T. brucei*, proteínas contendo motivos PPR são relatadas com

funções no processamento 5’ de tRNA (20) e também na poliadenilação, poliuridinação e tradução (21).

Para cada uma das proteínas selecionadas neste trabalho foi possível avançar no que concerne a confirmação de sua localização subcelular e ainda, para algumas delas, quanto a sua importância para a viabilidade de células procíclicas de *T. brucei*. Os resultados indicam comportamentos distintos para cada uma delas na sua resposta a indução de RNAi, sugerindo possíveis diferenças funcionais. Novas abordagens, como a purificação e caracterização de complexos proteicos associados, fazem-se necessárias para melhor compreender a função de cada um e seu papel ou não no controle do metabolismo e tradução de mRNAs codificantes de proteínas ribossomais.

REFERÊNCIAS BIBLIOGRÁFICAS

1. Clayton,C. and Shapira,M. (2007) Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. *Mol Biochem Parasitol*, **156**, 93–101.
2. Kramer,S. and Carrington,M. (2011) Trans-acting proteins regulating mRNA maturation, stability and translation in trypanosomatids. *Trends Parasitol.*, **27**, 23–30.
3. Clayton,C. (2013) The Regulation of Trypanosome Gene Expression by RNA-Binding Proteins. *PLoS Pathog.*, **9**, 9–12.
4. Erben,E.D., Fadda,A., Lueong,S., Hoheisel,J.D. and Clayton,C. (2014) A Genome-Wide Tethering Screen Reveals Novel Potential Post-Transcriptional Regulators in *Trypanosoma brucei*. *PLoS Pathog.*, **10**.
5. Lueong,S., Merce,C., Fischer,B., Hoheisel,J.D. and Erben,E.D. (2016) Gene expression regulatory networks in *Trypanosoma brucei*: Insights into the role of the mRNA-binding proteome. *Mol. Microbiol.*, **100**, 457–471.
6. da Costa Lima,T.D., Moura,D.M.N., Reis,C.R.S., Vasconcelos,J.R.C., Ellis,L., Carrington,M., Figueiredo,R.C.B.Q. and de Melo Neto,O.P. (2010) Functional characterization of three Leishmania poly(A) binding protein homologues with distinct binding properties to RNA and protein partners. *Eukaryot. Cell*, **9**, 1484–1494.
7. Moura,D.M., Reis,C.R., Xavier,C.C., da Costa Lima,T.D., Lima,R.P., Carrington,M. and de Melo Neto,O.P. (2015) Two related trypanosomatid eIF4G homologues have functional differences compatible with distinct roles during translation initiation. *RNA*

Biol., **12**, 305–19.

8. Rosenzweig,D., Smith,D., Opperdoes,F., Stern,S., Olafson,R.W. and Zilberstein,D. (2008) Retooling Leishmania metabolism: from sand fly gut to human macrophage. *FASEB J.*, **22**, 590–602.

9. Robinson,K.A. and Beverley,S.M. (2003) Improvements in transfection efficiency and tests of RNA interference (RNAi) approaches in the protozoan parasite Leishmania. *Mol. Biochem. Parasitol.*, **128**, 217–28.

10. Sunter,J., Wickstead,B., Gull,K. and Carrington,M. (2012) A new generation of T7 RNA polymerase-independent inducible expression plasmids for *Trypanosoma brucei*. *PLoS One*, **7**.

11. Anderson,C.W., Baum,P.R. and Gesteland,R.F. (1973) Processing of adenovirus 2-induced proteins. *J. Virol.*, **12**, 241–252.

12. Zoltner,M., Krienitz,N., Field,M.C. and Kramer,S. (2018) Comparative proteomics of the two *T. brucei* PABPs suggests that PABP2 controls bulk mRNA. *PLoS Negl. Trop. Dis.*, **12**, 1–18.

13. Wurst,M., Robles,A., Po,J., Luu,V.D., Brems,S., Marentije,M., Stoitsova,S., Quijada,L., Hoheisel,J., Stewart,M., *et al.* (2009) An RNAi screen of the RRM-domain proteins of *Trypanosoma brucei*. *Mol. Biochem. Parasitol.*, **163**, 61–65.

14. Alsford,S., Turner,D.J., Obado,S.O., Sanchez-flores,A., Glover,L., Berriman,M., Hertz-fowler,C. and Horn,D. (2011) High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. *Genome Res.*, **21**, 915–924.

15. Kelley,L., Mezulis,S., Yates,C., Wass,M. and Sternberg,M. (2015) The Phyre2 web portal for protein modelling, prediction, and analysis. *Nat. Protoc.*, **10**, 845–858.

16. Lee,E.K. and Diehl,J.A. (2014) SCFs in the new millennium. *Oncogene*, **33**, 2011–2018.

17. Urbaniak,M.D., Martin,D.M.A. and Ferguson,M.A.J. (2013) Global quantitative SILAC phosphoproteomics reveals differential phosphorylation is widespread between the procyclic and bloodstream form lifecycle stages of *Trypanosoma brucei*. *J. Proteome Res.*, **12**, 2233–2244.

18. Dean,S., Sunter,J.D. and Wheeler,R.J. (2017) TrypTag.org: A Trypanosome Genome-wide Protein Localisation Resource. *Trends Parasitol.*, **33**, 80–82.

19. Pusnik,M., Small,I., Read,L.K., Fabbro,T. and Schneider,A. (2007)

Pentatricopeptide Repeat Proteins in *Trypanosoma brucei* Function in Mitochondrial Ribosomes. *Mol. Cell. Biol.*, **27**, 6876–6888.

20. Taschner,A., Weber,C., Buzet,A., Hartmann Roland K.,R.K., Hartig,A. and Rossmanith,W. (2012) Nuclear RNase P of *Trypanosoma brucei*: A Single Protein in Place of the Multicomponent RNA-Protein Complex. *Cell Rep.*, **2**, 19–25.

21. Aphasizheva,I., Maslov,D., Wang,X., Huang,L. and Aphasizhev,R. (2011) Pentatricopeptide Repeat Proteins Stimulate mRNA Adenylation/Uridylation to Activate Mitochondrial Translation in Trypanosomes. *Mol. Cell*, **42**, 106–117.

Tabela 1. Iniciadores utilizados para amplificação da ORF do gene alvo e para construção do cassete para marcação do gene endógeno com eYFP. As bases em negrito representam os sítios de restrição BamHI (GGATCC), BglII (AGATCT) ou HindIII (AAGCTT); F (*Foward*) e R (*Reverse*).

Para amplificar a ORF 5' → 3'	
RBP23 F	ATGGT**GGATCC**TTCTGTGCTGCTGGC
RBP23 R	**AAGCTT**TCAGTTGTGGACTTGCACCC
Tb927.10.14700 F	**AAGCTT**ATGAACGCGAGAGGT
Tb927.10.14700 R	**GGATCC**TTACCATACACCCTCACCGTT
Tb927.10.810 F	**AAGCTT**ATGGGGGCGGAAAAA
Tb927.10.810 R	**AGATCT**TTAATTATCACGCTTCTTTTT
Tb927.7.7.7460 F	**AAGCTT**ATGGTTTTTTCCACACACGG
Tb927.7.7.7460 R	**GGATCC**TTACCTCCTCACCGTTGGTTT
Tb927.9.7080 F	**AAGCTT**ATGCCGGATTACACG
Tb927.9.7080 R	**GGATCC**TCACACGGGCCGACCGCGCCC
Para marcação com eYFP 5' → 3'	
RBP23 F	GGAAAATTATCACCGTTAGCAGTGGAGCAATAACAATCCCGGGACTATTAG CAGCAGTGGTGGGGGTGCAAGTCCACAACGGATCCGGTGGCGGAATGGT
RBP23 R	AGCAAACAGGTCAAAGGATCATGACGACTTCTCTGGTACCGTGCTTCGCTT TTTATTCGACTACTTAATTTATCTAGCTTGAAGAACTCGTCAAGAAGGC
Tb927.10.14700 F	GTGCCTCTAAAGGGTTTTTGTTATTTTTCTGCTAGAAAAGGATTGTCTGTAG TAATTTTTATTTTGTCGGTTGATTGTTAATGCCTTTGTCTCAAGAAGA
Tb927.10.14700 R	TTTTCCTCTTTATTTGGCCGACTGCGCGTACCATCCTTCTCATCCGCCTTCA CGGGTCGTCCTTTACCTCTCGCGTTCATGGTGGCGGAAGCTTGAGAAC
Tb927.10.810 F	CTCTCCTGTCGAAAATACTTTATATACTTCTGTCTGCACCAGTTGTGTCAGT GTTCTGTTCTTTAGTCTGAAGGTCATGAATGCCTTTGTCTCAAGAAGA
Tb927.10.810 R	ATCACCTCTGTAAATGCGCCAACATTCTCTGCAACCTCACCAGAGTTGGTG GGACGATTTATTTTTTTTTCCGCCCCCATGGTGGCGGAAGCTTGAGAAC
Tb927.7.7.7460 F	AAATCGGTCCAAGTGATCTTGATGAAGCTGCATGTCGATATACGTGGCTTG ACCAAGCTATCAAACCAACGGTGAGGAGGGGATCCGGTGGCGGAATGGT
Tb927.7.7.7460 R	ACAGAGCTATTAAAGGTGTGGAGTGAAGGACCCACTTTACTTTTCTTTTTTG CTTTCTCATCTCCCCTTTCCTTGTTCATGAAGAACTCGTCAAGAAGGC
Tb927.9.7080 F	TCTACGCAGTGTGAAAGGTTTCACAACACTAGCATCATCGTCGCTTTTATTA TTAATACTATCACAGGTGAAGCTCACCGATGCCTTTGTCTCAAGAAGA
Tb927.9.7080 R	CGACTATACGCAGAGGTCTCACTTGGTGGGATGTTCGGGTACTCGTTAAA GTACTTGATGAACTGCGTGTAATCCGGCATGGTGGCGGAAGCTTGAGAAC

Tabela 2. Antibióticos e suas respectivas concentrações para manutenção das células transfectadas.

Plasmídeo/Cassete	Antibiótico	Procíclica	Sanguínea
pSMOX	Puromicina	1 µg/ml	0.1 µg/ml
p2T7177	Fleomicina	2.5 µg/ml	-
p3666	Blasticidina	10 µg/ml	5 µg/ml
N-terminal	Blasticidina	10 µg/ml	-
C-terminal	G418	15 µg/ml	-

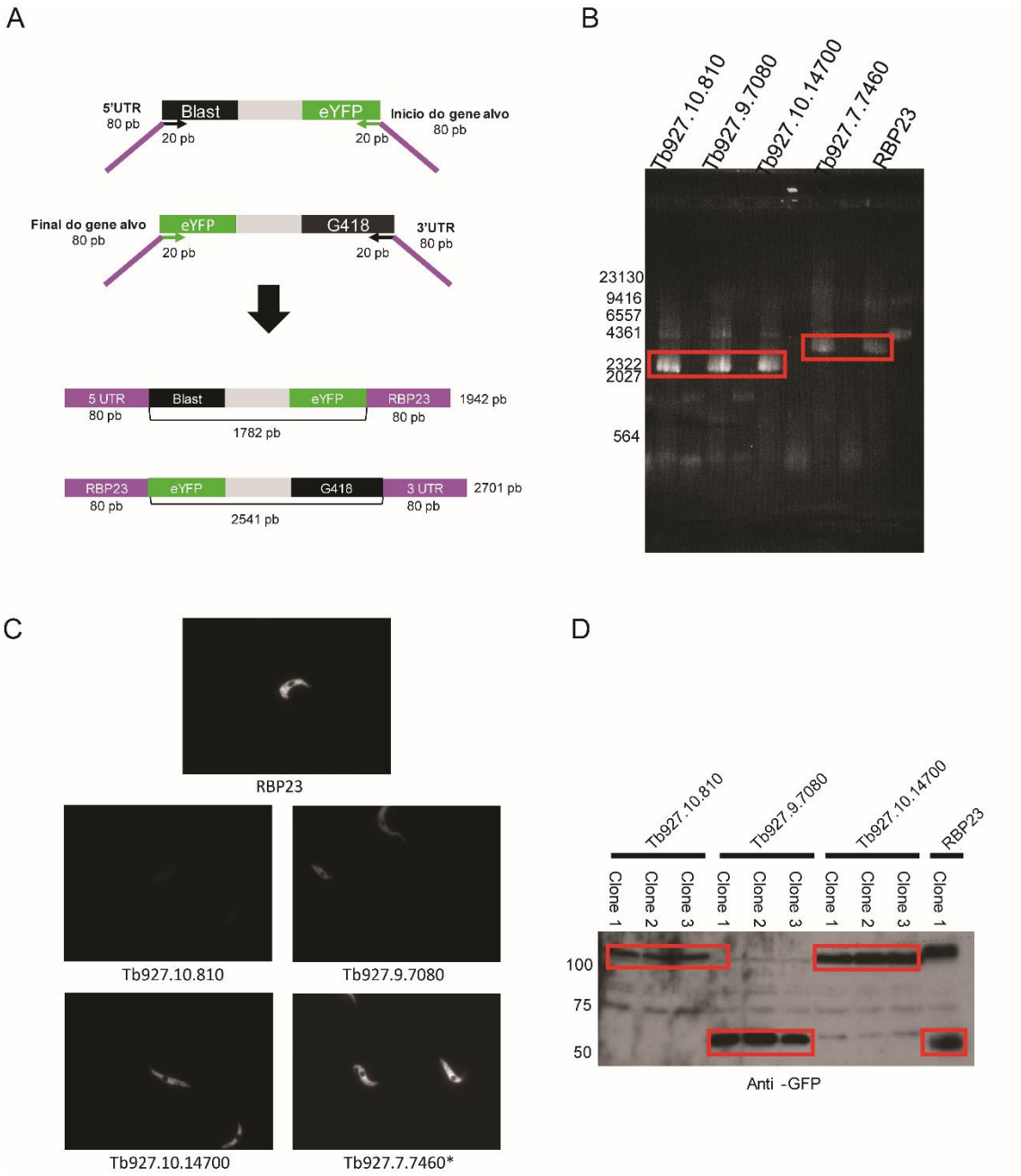


Figura 1. Expressão da RBP23 e proteínas parceiras marcadas com eYFP. A) Esquema dos iniciadores utilizados para amplificação dos fragmentos para marcação N- ou C-terminal. **B)** Gel de agarose 1% mostrando os fragmentos amplificados nos seus respectivos tamanhos de 1782 pb para marcação N-terminal e 2541 pb para C-terminal. **C)** Microscopia de cinco linhagens celulares expressando cada uma delas, uma proteína marcada com eYFP (*) única linhagem celular na qual houve falhas na transfecção seguinte com o plasmídeo de RNAi. Imagens na objetiva de 100x; **D)** *Western blot* com anti-GFP de extratos celulares de cada linhagem celular expressando eYFP. Três clones foram utilizados de cada com exceção para a RBP23 no qual só foi possível recuperar um clone.

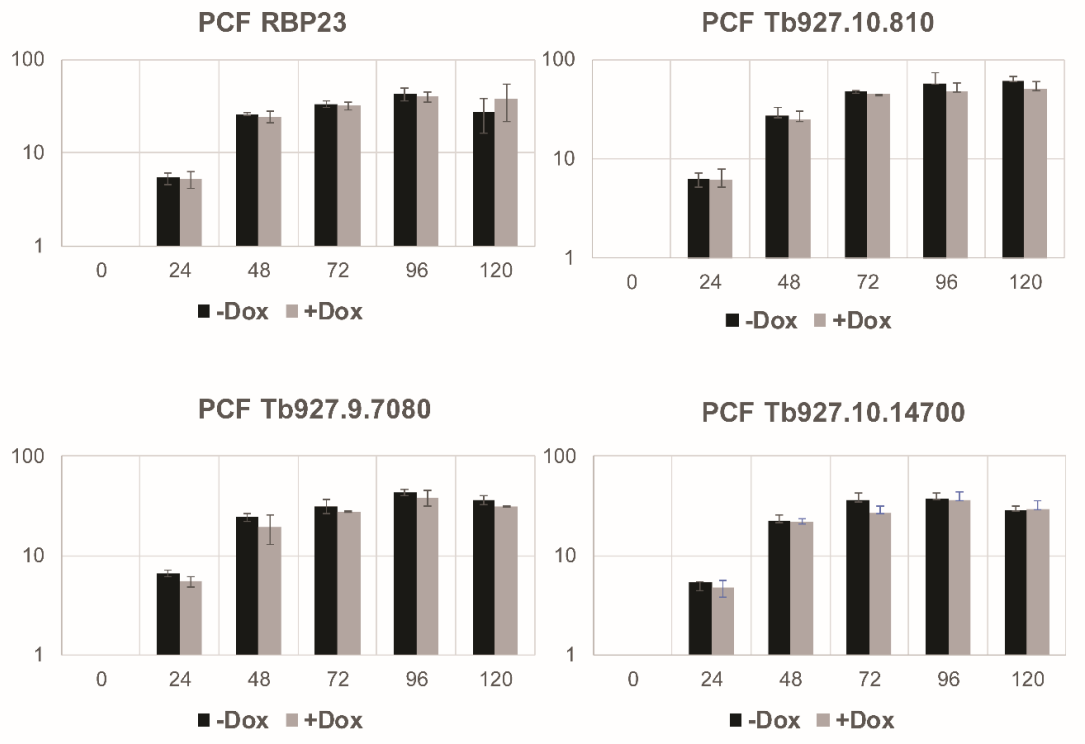


Figura 2. Curva de crescimento após indução de RNAi da RBP23 e proteínas parceiras em *T. brucei*. Nos gráficos o eixo Y representa a contagem de células procíclicas (PCF), onde 1 equivale a 10^6 (escala logarítmica). O eixo X representa o tempo de indução de RNAi o qual foi de 120 horas em células sem (preto) e com (cinza) doxiciclina (Dox).

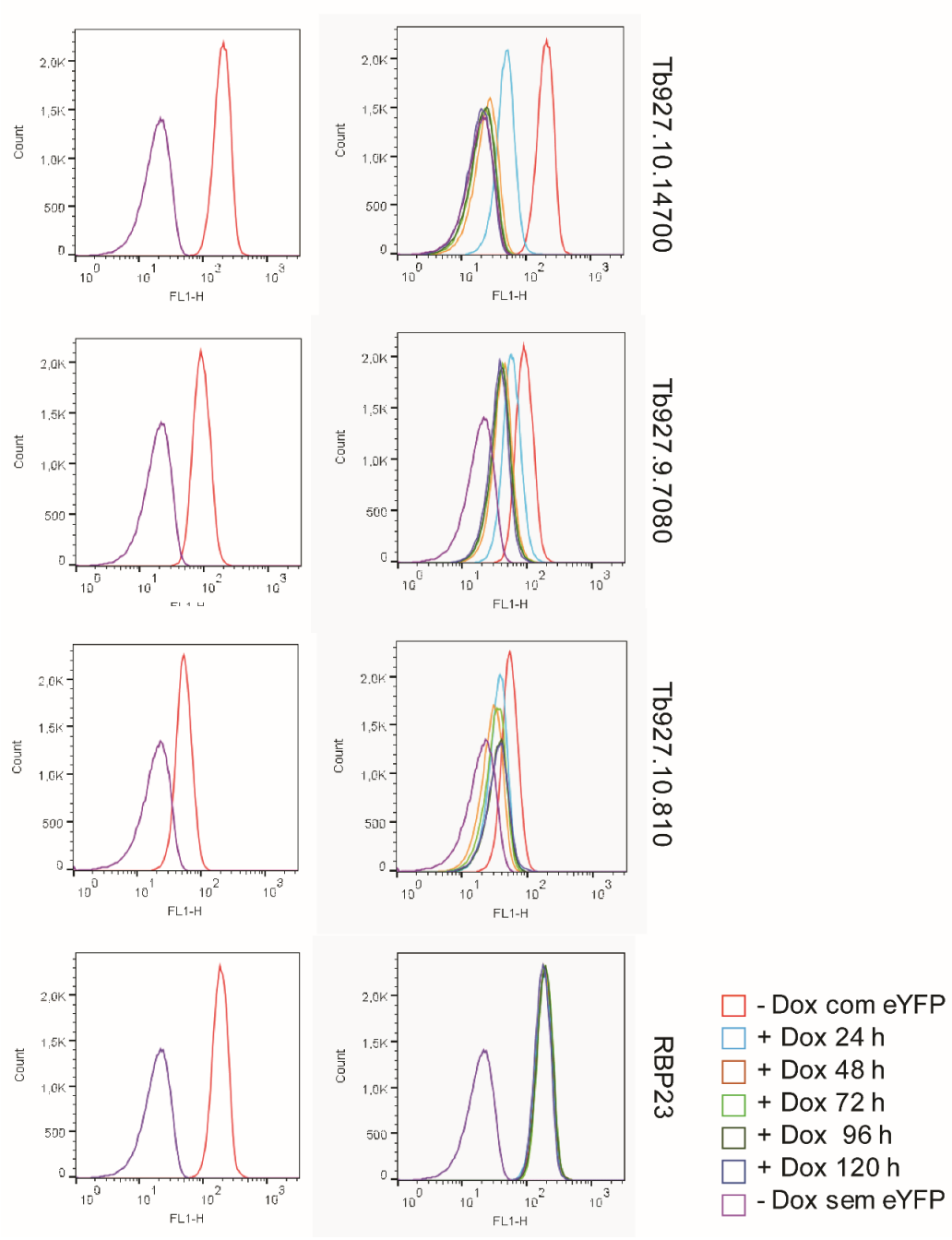


Figura 3. Medição da quantidade de fluorescência após indução de RNAi da RBP23 e proteínas parceiras em *T. brucei*. O eixo X representa a contagem de células enquanto o eixo representa a quantidade de fluorescência. O pico roxo representa células sem fluorescência enquanto o pico vermelho representa células fluorescentes ambas sem adição de doxiciclina (-Dox). Os picos tendem a mover da direita para a esquerda após adição de doxiciclina (+Dox) por consequência da perda da fluorescência.

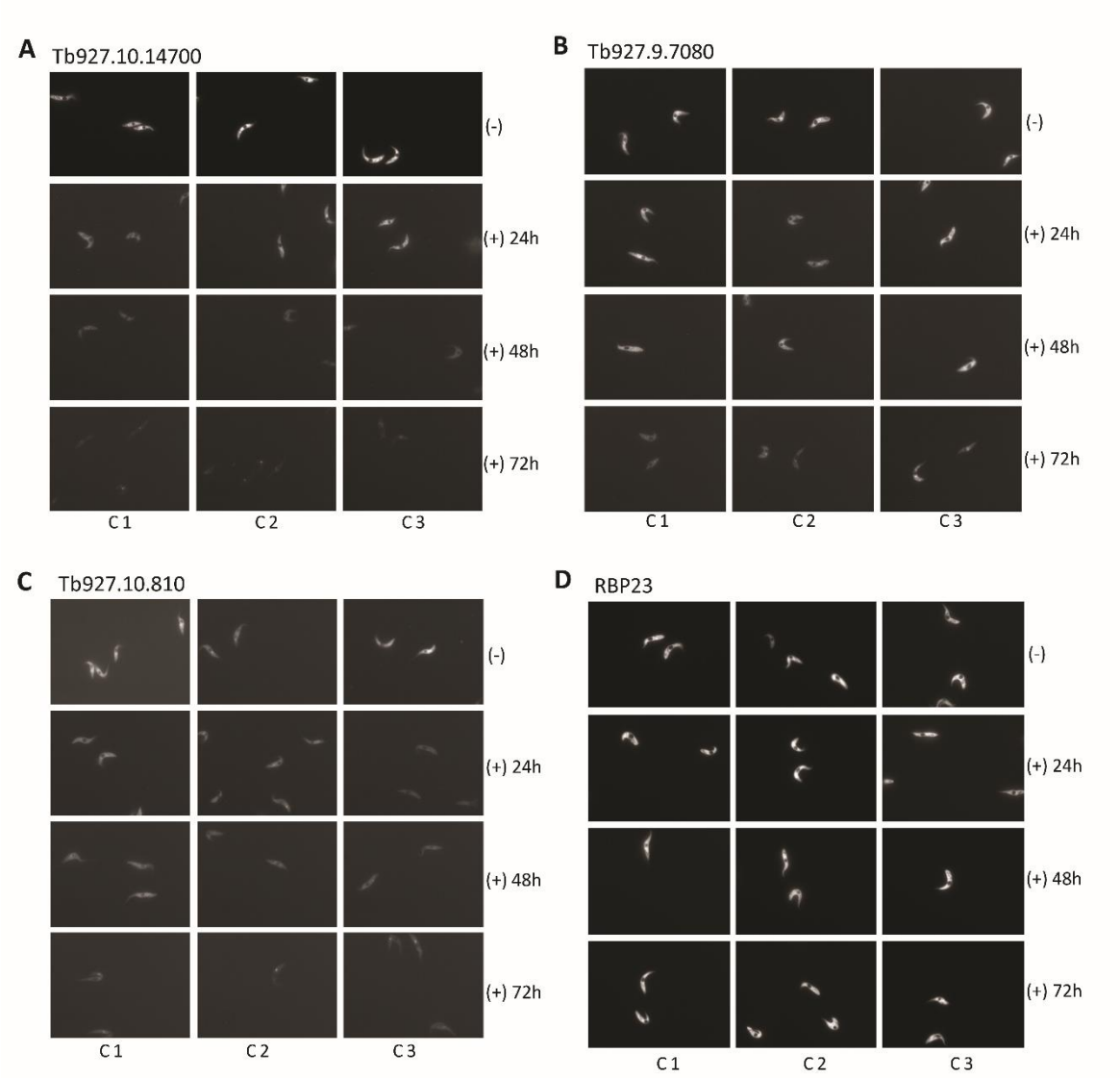


Figura 4. Microscopia de três clones após indução de RNAi da RBP23 e proteínas parceiras em *T. brucei*. As células de três clones (C1, C2 e C3) foram monitoradas por três dias após adição (+) de doxiciclina. Utilizado exposição fixa de 4000ms do filtro de YFP e imagens na objetiva de 100x.

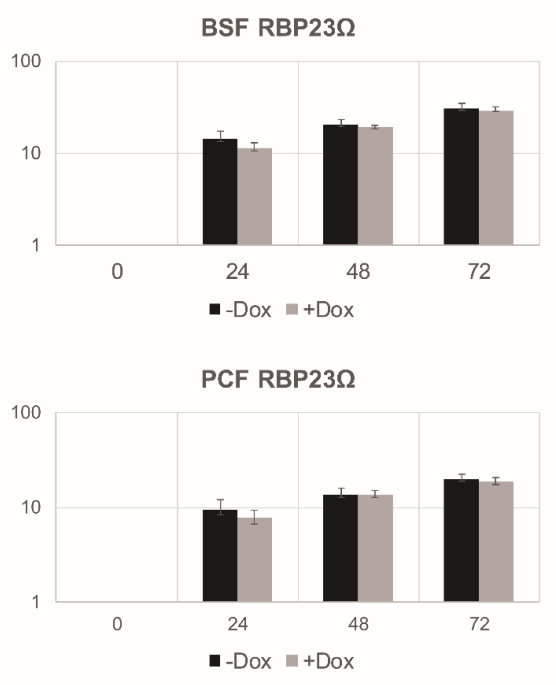


Figura 5. Curva de crescimento após RNAi da RBP23 em *T. brucei*. O experimento foi realizado em células procíclicas (PCF) e em sanguíneas (BSF) durante três dias em células sem (preto) e com (cinza) doxiciclina (Dox). Para BSF, o número 1 do eixo Y equivale a 10^5 e em PCF 1 corresponde a 10^6. Ω representa o RNAi em forma de grampo.

6 DISCUSSÃO GERAL

Em diversos organismos, as proteínas com domínios de ligação ao RNA (RBPs) tem sido envolvidas em importantes funções ligadas a regulação da expressão gênica e ao metabolismo dos mRNAs, participando em processos que envolvem desde a síntese até a degradação dos mesmos (MARONDEDZE *et al*., 2016). Na presente tese, três RBPs de *Leishmania infantum* foram selecionadas, RBP23, DBRD2 e ZC3H41 para uma melhor caracterização de suas interações com outras proteínas de ligação ao RNA como as PABPS e também com seus mRNAs alvos.

Em *L. infantum*, A RBP23 e ZC3H41 foram co-imunoprecipitadas com a PABP1 (DE MELO NETO *et al*., 2018), a qual interage com os fatores de iniciação da tradução EIF4E4 e EIF4G3. Todas essas proteínas, com exceção da ZC3H41, foram definidas como ativadores da tradução em *T. brucei*, o que sugere uma função em comum durante a iniciação da tradução (ERBEN *et al*., 2014). A RBP23 interage diretamente com a PABP1 e com as outras duas PABPs pois todas possuem presumivelmente o possível domínio de interação com a RBP23, mapeado a princípio no RRM4 na região N-terminal. Em humanos, os RRM3 e 4 da PABP são capazes de se ligar com baixa afinidade as sequencias poli(A) dos mRNAs. Esses RRMs também interagem com elementos ricos em AU do RNA e são capazes de mediar interações proteína-proteína (GOSS; KLEIMAN, 2013). Nossos dados sugerem que, embora as três PABPs sejam capazes de se ligar a RBP23 (e também a DRBD2) *in vitro*, outras interações que ocorrem *in vivo*, devam permitir que apenas a interação direta entre PABP1 e RBP23 se mantenha. De fato, outras proteínas foram encontradas co-imunoprecipitadas com PABP1, RBP23 e ZC3H41: a LinJ.18.0300 e a LinJ.05.0450 com funções ainda não definidas e que talvez contribuam para essa especificidade (Figura 12). A LinJ.18.0300 contém os domínios NTF2 (*Nuclear Transport Factor* 2) e RRM, presentes em proteínas ativadoras de GTPase que podem estimular o transporte de proteínas do citoplasma para o núcleo (LUI, 2009). Já a LinJ.05.0450 é estruturalmente similar a proteína SKP1 um componente do complexo ubiquitina ligase que sinaliza através da ubiquitinilação de proteínas alvos a degradação pelo proteossoma (SAKAMOTO et al., 2001). Em *T. brucei* as cinco proteínas mais enriquecidas co-precipitadas com a

PABP1 são EIF4G3, EIF4E4, PABP1, RBP23, ZC3H41 e a proteína hipotética Tb927.7.7460 (homólogo da LinJ.05.0450) (ZOLTNER *et al*., 2018). Estes resultados confirmam a forte associação ao menos entre a LinJ.05.0450 com a RBP23 e proteínas parceiras, sugerindo uma função crítica, talvez mediando a ligação específica da RBP23 a PABP1.

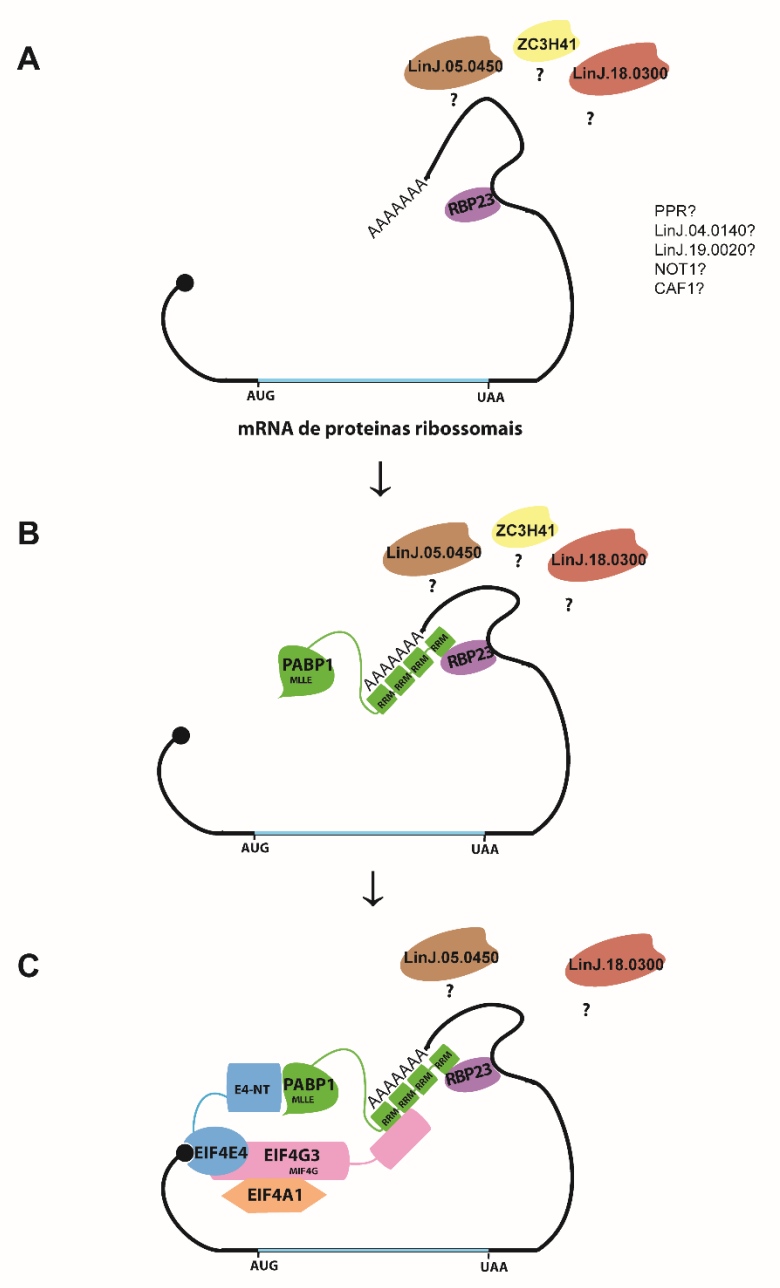


Figura 10. Modelo proposto para a RBP23. A RBP23 interage com o motivo 3'UTR no mRNA da proteína ribossomais (A) e, em seguida, interage com a PABP1 provavelmente via RRM4 (B), recrutando o complexo EIF4E4 / EIF4G3 / EIF4AI (C). Em algum momento é possível que a proteína ZC3H41 se dissocie do complexo, uma vez que esta não parece precipitar eficientemente com o complexo EIF4E4 / EIF4G3 / EIF4AI.

A proteína ZC3H41 tem uma expressão constitutiva e não parece sofrer modificações pós-traducionais. Esta proteína contém três classes de domínios diferentes, no qual se destaca na região C-terminal o domínio helicase, similar as proteínas DHH1, eIF4AI e eIF4AIII. O silenciamento da proteína ZC3H41 mostrou que essa é essencial para a sobrevivência do *T. brucei* (ELIAZ *et al*., 2017). De forma similar, a helicase EIF4AI, que também tem expressão constitutiva durante o ciclo de vida do parasita também é essencial para a síntese proteica e para a viabilidade celular (DHALIA *et al*., 2006). Em *T. brucei*, os RNAs de sequência líder interagem com um conjunto de proteínas, dentre elas, a ZC3H41 (ELIAZ *et al*., 2017) e apesar de distribuída no citoplasma migra para os grânulos sob condições de estresse (DEAN *et al*., 2017; ELIAZ *et al*., 2017). Além de interagir como a PABP1 e com um conjunto específico de proteínas parceiras que serão melhor discutidas adiante, essas proteínas também são capazes de interagir com a PABP2 e PABP3. A ZC3H41 pode dessa forma participar de dois complexos distintos, um com a PABP1 atuando na regulação da tradução, e outro com a PABP2-PABP3 que pode também atuar na tradução além de uma possível função nuclear de ambas as proteínas (Figura 11). É possível que em ambos os casos a ZC3H41 atue facilitando a atuação das moléculas de PABP, independente da sua interação com fatores de tradução.

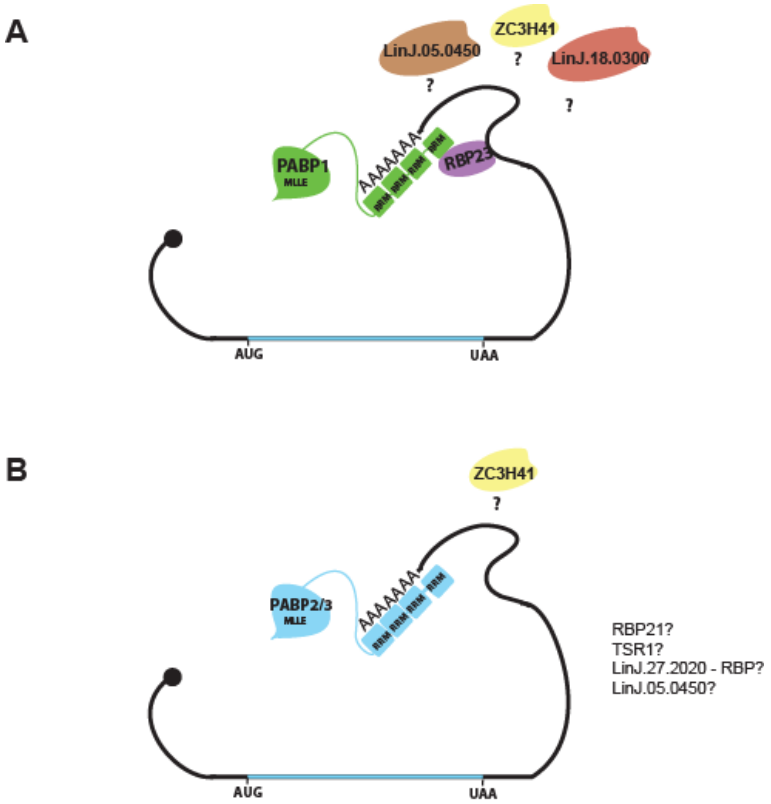


Figura 11. Modelo proposto para a ZC3H41. (A) A ZC3H41 interage com o complexo formado pela RBP23, LinJ.05.0450, LinJ.18.0300 e PABP1; (B) ou em um complexo independente formado pelas PABP2/3 e outras proteínas identificadas com a ZC3H41 talvez atuem nesse complexo.

A proteína DRBD2 foi selecionada para este trabalho por ter sido co-imunoprecipitada especificamente com a PABP2 em ensaios preliminares de *L. infantum* (Manuscrito II – ASSIS *et al*., 2019), dado confirmado mais recentemente com seus homólogos de *T. brucei*. (ZOLTNER *et al*., 2018). Nesta tese, esta proteína foi encontrada associada a múltiplas proteínas contendo diferentes domínios de ligação ao RNA, do tipo CCCH, Pumílio (PUF) e um ou mais RRMs. A PUF6 é um repressor da tradução em *Saccharomyces cerevisiae* (SHAHBABIAN *et al*., 2014) e foi envolvida no decaimento de mRNA mediado por retroposon de *Leishmania* (AZIZI *et al*., 2017). Em *T. brucei* a DRBD2, e os ortólogos de proteínas parceiras ZC3H39, ZC3H40, RBP12 e a proteína hipotética Tb927.1.670 são descritas como repressoras da tradução, (LUEONG *et al*., 2016). Porém a interação com a PABP2, uma ativadora da tradução, pode indicar que a DRBD2 atue em complexos distintos com funções diferentes (Figura 12). Um número considerável de proteínas mitocondriais também foram encontradas associadas à DRBD2, sugerindo uma possível função mitocondrial. A interação mais importante, porém, parece ser

entre a DRBD2 e as já citadas PABP2 e PABP3. Estes dois homólogos da PABP interagem entre si e também com os mesmos mRNAs (Manuscrito II – Assis *et al*., 2019; DA COSTA LIMA *et al*., 2010) e ainda migram para o núcleo quando a transcrição é inibida por Actinomicina D (KRAMER *et al.,* 2013; DA COSTA LIMA *et al*., 2010)**,** sugerindo uma função nuclear quando não ocorre transcrição e consequentemente tradução. De forma semelhante a DRBD3 uma proteína de estabilização que em *T. brucei* é normalmente encontrada no citoplasma, mas sob estresse oxidativo continua ligada ao mRNA, mas migra para o núcleo. Um típico comportamento de proteínas que transportam mRNAs específicos (FERNÁNDEZ-MOYA *et al*., 2012). A PABP2 também está associada a tradução, apesar de apresentar várias substituições nos resíduos envolvidos no reconhecimento da cauda poliA, o que indica que reconhece mRNAs mais específicos (GUERRA *et al*., 2011; DA COSTA LIMA *et al*., 2010). É esperado que os transcritos sejam mais diferentes dos encontrados com a PABP1 e mais similares aos encontrados com proteínas parceiras, como a DRBD2, como de fato foi observado.

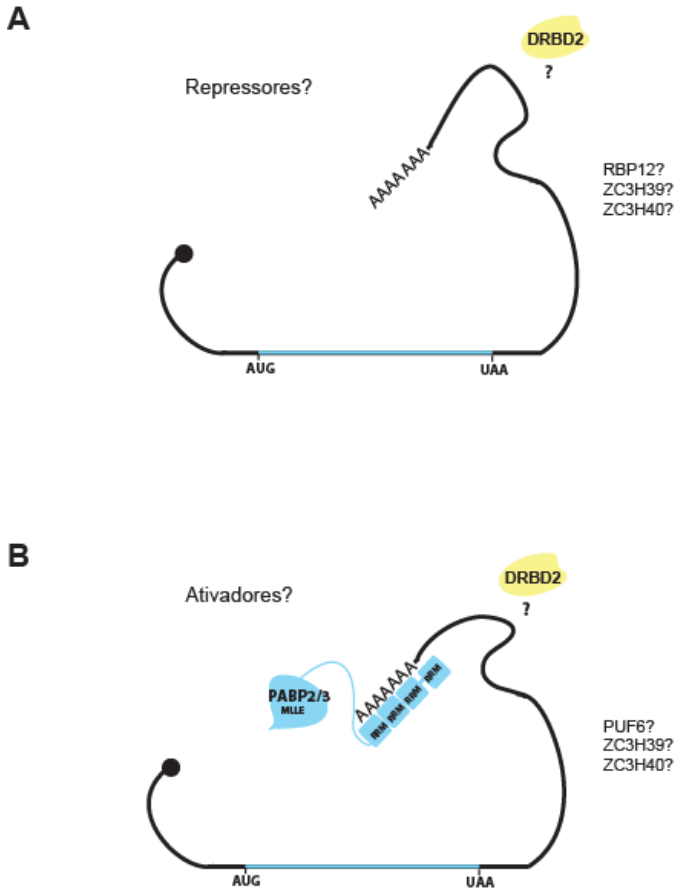


Figura 12. Modelo proposto para a DRBD2. (A) A DRBD2 atua em complexos distintos possivelmente reprimindo a tradução como as proteínas RBP12, ZC3H39 e ZC3H40 na ausência da PABP2/3; (B) ou ativando a tradução com as proteínas PUF6, ZC3H39 e ZC3H40 na presença da PABP2/3.

A grande porcentagem dos transcritos encontrados co-imunoprecipitadas com a PABP1, a RBP23 e a ZC3H41 são mRNAs de proteínas ribossomais, diferentemente do que ocorre para a DRBD2, que possui um perfil bem mais variado, validando o resultado para as outras três proteínas. Os mRNAs de proteínas ribossomais de mamíferos compartilham na 5' UTR sequencias ricas em pirimidinas conhecidos como elementos TOP, responsáveis pela regulação da sua tradução (HAMILTON *et al*., 2006; LEDDA *et al*., 2005). A proteína LARP1 associada ao motivo TOP age como um repressor da tradução. Essa proteína também interage com a PABP formando um mRNA de conformação fechada inativo. A LARP1 é liberada do motivo TOP através de fosforilação pela mTORC1. Essa quinase também fosforila a 4E-BP1 um outro repressor da tradução, permitindo que o eIF4G interaja com o eIF4E (PHILIPPE *et al*., 2018; FONSECA *et al*., 2015). Entretanto, os tripanossomatídeos não possuem a região 5' UTR enriquecida de polipirimidinas,

então, a regulação deve ser diferente e talvez localizada na 3’ UTR dos transcritos (JENSEN *et al*., 2014). De fato, a interação direta entre PABP1 e EIF4E4 pode levar a uma conformação fechada dos mRNAs codificantes de proteínas ribossomais semelhante a mediada pela proteína LARP, e coincidentemente todos são fosforilados em condições equivalentes (DE MELO NETO *et al*., 2018). Nessa condição não seria necessária uma proteína com motivos para a LARP na 5’ UTR, mas sim uma outra proteína que dê especificidade aos mRNAs em questão, no caso talvez a RBP23. O modelo proposto baseado em nossos dados para o complexo sugere que elementos ricos em pirimidinas presentes na região 3’ UTR do mRNA de proteínas ribossomais são o alvo da RBP23 que interage com a PABP1 já ligada a cauda poli-A do mRNA. A interação é um sinal e age como uma ponte para que o complexo EIF4E4/EIF4G3 se ligue ao *cap* desses mRNAs específicos. Curiosamente a RBP23 interage com o próprio transcrito. Em *S. cerevisiae*, a 4E-BP proteína repressora da tradução conhecida com Cf20p também interage com o próprio transcrito e pode representar um mecanismo auto regulatório (COSTELLO *et al*., 2015). Como previamente observado em estudos *in vitro* com a PABP de humanos, nos quais a ligação da PABP ao trato adenilado presente na 5 'UTR do seu próprio mRNA é necessária e suficiente para reprimir a tradução deste (DE MELO NETO *et al*., 1995), é possível que a ligação da RBP23 ao seu mRNA também de alguma forma esteja envolvida em um evento de autoregulação associada ao controle da sua própria síntese.

Esta tese gerou contribuições importantes sobre a função de diferentes RBPs e mecanismos de controle da tradução de mRNAs específicos de tripanosomatídeos. Mais estudos serão necessários, contudo, para avaliar se existe uma interação direta entre a ZC3H41 e homólogos de PABP ou com a própria RBP23. Da mesma forma é preciso esclarecer se as proteínas LinJ.05.0450 e LinJ.18.0030, ou seus ortólogos de *T. brucei* formam de fato o denominado complexo RBP23, proposto aqui, através de interações diretas entre estas proteínas. Um segundo complexo formado pela ZC3H41 com a PABP2 e PABP3 também foi sugerido e precisa de uma melhor caracterização assim como é necessário se investigar se a DRBD2 auxilia realmente na função da PABP2. Novos ensaios serão realizados com o anti-ZC3H41, o qual foi comprovado ser eficiente e especifico, com o intuito de verificar se a proteína endógena também possui o mesmo perfil. Como

não há dados de localização da ZC3H41 em *Leishmania*, o anticorpo também será usado para localização subcelular da proteína sem e sob a ação de diferentes inibidores. Outro passo a se investigar é o comportamento da DRBD2 na presença de inibidores de transcrição, uma vez que em *T. brucei* essa proteína é encontrada acumulada na região ao redor do núcleo (DEAN *et al*., 2017). No conjunto os resultados alcançados e os que devem ser obtidos com os experimentos propostos devem ajudar a melhor delinear aspectos importantes do controle da expressão gênica e da síntese proteica em protozoários tão importantes no contexto da saúde pública mundial.

7 CONCLUSÕES

1. A PABP1 interage com um conjunto específicos de proteínas hipotéticas, com o complexo EIF4G3/EIF4E4 e com a proteína de ligação ao RNA, RBP23;

2. A RBP23 interage com o complexo de iniciação da tradução EIF4G3/EIF4E4/PABP1, e diretamente com a PABP1. Esse complexo tem uma preferência de ligação a mRNAs de proteínas ribossomais. Já a DRBD2, uma segunda RBP investigada aqui, interage preferencialmente com as PABP2/PABP3, e com um perfil de mRNAs mais diversos e distintos dos encontrados com as RBP23/PABP1.

3. A proteína ZC3H41 interage com o complexo da RBP23 que inclui a PABP1, porém também é capaz de interagir independentemente com as PABP2 e PABP3. Similar a RBP23 e a PABP1, a ZC3H41 tem uma preferência por mRNAs de proteínas ribossomais;

4. A caracterização funcional de proteínas do complexo RBP23 em linhagens celulares de *T. brucei,* indicaram que a princípio pelo menos uma das proteínas parceiras da RBP23, a Tb927.10.14700, não é essencial para a viabilidade celular.

5. A RBP23, PABP1 e ZC3H41 fazem parte do mesmo complexo que interage principalmente com mRNAs de proteínas ribossomais, denominado complexo RBP23, que pode fazer a seleção dos mRNAs que serão traduzidos pelo complexo EIF4E4/EIF4G3.

REFERÊNCIAS

AKHOUNDI, M. et al. A Historical Overview of the Classification, Evolution, and Dispersion of Leishmania Parasites and Sandflies. **PLoS Neglected Tropical Diseases**, v. 10, n. 3, p. 1–40, 2016.

AZIZI, H.; DUMAS, C.; PAPADOPOULOU, B. The Pumilio-domain protein PUF6 contributes to SIDER2 retroposon-mediated mRNA decay in Leishmania. **Rna**, v. 23, n. 12, p. 1874–1885, 2017.

BAKER, C. H.; WELBURN, S. C. The Long Wait for a New Drug for Human African Trypanosomiasis. **Trends in Parasitology**, v. 34, n. 10, p. 818–827, 2018.

BARRETT P., M. et al. The trypanosomiases. **Lancet**, v. 362, n. 9394, p. 1469–1480, 2003.

BATES, E. J.; KNUEPFER, E.; SMITH, D. F. Poly (A)-binding protein I of Leishmania : functional analysis and localisation in trypanosomatid parasites. **Nucleic Acids Research**, v. 28, n. 5, p. 1211–1220, 2000.

BATES, P. A. Transmission of Leishmania metacyclic promastigotes by phlebotomine sand flies. **International Journal for Parasitology**, v. 37, n. 10, p. 1097–1106, 2007.

BENZ, C. et al. Messenger RNA processing sites in Trypanosoma brucei. **Molecular and Biochemical Parasitology**, v. 143, n. 2, p. 125–134, 2005.

BERRIMAN, M.; GHEDIN, E.; HERTZ-FOWLER, C. The genome of the African trypanosome, Trypanosoma brucei. **Science**, v. 309, p. 416–422, 2005.

BESTEIRO, S. et al. Protein turnover and differentiation in Leishmania. **International Journal for Parasitology**, v. 37, n. 10, p. 1063–1075, ago. 2007.

BLUM, J.; NEUMAYR, A.; LOCKWOOD, D. Treatment of Tegumentary Forms of Leishmaniasis. In: **The Leishmaniases: Old Neglected Tropical Diseases**. p. 191–225, 2018.

BRASIL, M. DA S. **Leishmaniose Visceral: o que é, causas, sintomas, tratamento, diagnóstico e prevenção**. Disponível em: <http://portalms.saude.gov.br>.

BRASIL, M. DA S. **Leishmaniose Tegumentar (LT): o que é, causas, sintomas, tratamento, diagnóstico e prevenção**. Disponível em: <http://portalms.saude.gov.br>.

CASSOLA, A.; DE GAUDENZI, J. G.; FRASCH, A. C. Recruitment of mRNAs to cytoplasmic ribonucleoprotein granules in trypanosomes. **Molecular Microbiology**, v. 65, n. 3, p. 655–670, 2007.

CAVALCANTI, D. P.; DE SOUZA, W. The Kinetoplast of Trypanosomatids: From Early Studies of Electron Microscopy to Recent Advances in Atomic Force Microscopy. **Scanning**, v. 2018, p. 1–10, 2018.

CAVALIER-SMITH, T. Higher classification and phylogeny of Euglenozoa. **European Journal of Protistology**, v. 56, p. 250–276, 2016.

CECCHI, G. et al. Developing a continental atlas of the distribution and trypanosomal infection of tsetse flies (Glossina species). **Parasites and Vectors**, v. 8, n. 1, 2015.

CHENG, S.; GALLIE, D. R. eIF4G, eIFiso4G, and eIF4B bind the poly(A)-binding protein through overlapping sites within the RNA recognition motif domains. **Journal of Biological Chemistry**, v. 282, n. 35, p. 25247–25258, 2007.

CLAYTON, C. The Regulation of Trypanosome Gene Expression by RNA-Binding Proteins. **PLoS Pathogens**, v. 9, n. 11, p. 9–12, 2013.

CLAYTON, C.; SHAPIRA, M. Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. **Mol Biochem Parasitol**, v. 156, n. 2, p. 93–101, 2007.

COSTELLO, J. et al. Global mRNA selection mechanisms for translation initiation. **Genome Biology**, v. 16, n. 1, p. 1–21, 2015.

CRIBB, P.; SERRA, E. One- and two-hybrid analysis of the interactions between components of the Trypanosoma cruzi spliced leader RNA gene promoter binding complex. **International Journal for Parasitology**, v. 39, n. 5, p. 525–532, 2009.

DA COSTA LIMA, T. D. et al. Functional characterization of three Leishmania poly(A) binding protein homologues with distinct binding properties to RNA and protein partners. **Eukaryotic Cell**, v. 9, n. 10, p. 1484–1494, 2010.

DA COSTA LIMA, T. D. **Abordagens moleculares voltadas a caracterização funcional dos homólogos da proteína de ligação a cauda poli-A (PABP) em espécies de Leishmania sp .** Universidade Federal de Pernambuco, 2012.

DE MELO NETO, O. P. et al. The unique Leishmania EIF4E4 N-terminus is a target for multiple phosphorylation events and participates in critical interactions required for translation initiation. **RNA Biology**, v. 12, n. 11, p. 1209–1221, 2015.

DE MELO NETO, O. P. et al. Phosphorylation and interactions associated with the control of the Leishmania Poly-A Binding Protein 1 (PABP1) function during translation initiation. **RNA Biology**, v. 15, n. 6, p. 739–755, 2018.

DE MELO NETO, O. P.; STANDART, N.; DE SA, C. M. Autoregulation of poly(A)-binding protein synthesis in vitro. **Nucleic Acids Research**, v. 23, n. 12, p. 2198–2205, 1995.

DEAN, S.; SUNTER, J. D.; WHEELER, R. J. TrypTag.org: A Trypanosome Genome-wide Protein Localisation Resource. **Trends in Parasitology**, v. 33, n. 2, p. 80–82,

2017.

DHALIA, R. et al. Translation initiation in Leishmania major: Characterisation of multiple eIF4F subunit homologues. **Molecular and Biochemical Parasitology**, v. 140, n. 1, p. 23–41, 2005.

DHALIA, R. et al. The two eIF4A helicases in Trypanosoma brucei are functionally distinct. **Nucleic Acids Research**, v. 34, n. 9, p. 2495–2507, 2006.

DVORAK, V.; SHAW, J.; VOLF, P. Parasite Biology: The Vectors. In: **The Leishmaniases: Old Neglected Tropical Diseases**. p. 31–77, 2018.

EL-SAYED, N. M. et al. Comparative genomics of trypanosomatid parasitic protozoa. **Science (New York, N.Y.)**, v. 309, n. 5733, p. 404–409, 2005.

ELIAZ, D. et al. **Exosome secretion affects social motility in Trypanosoma brucei**. v. 13, n. 3, p. 1-38, 2017

ELISEEVA, I. A.; LYABIN, D. N.; OVCHINNIKOV, L. P. Poly(A)-binding proteins: structure, domain organization, and activity regulation. **Biochemistry. Biokhimiia**, v. 78, n. 13, p. 1377–91, 2013.

ERBEN, E. D. et al. A Genome-Wide Tethering Screen Reveals Novel Potential Post-Transcriptional Regulators in Trypanosoma brucei. **PLoS Pathogens**, v. 10, n. 6, 2014.

FERNÁNDEZ-MOYA, S. M. et al. Alterations in DRBD3 Ribonucleoprotein Complexes in Response to Stress in Trypanosoma brucei. **PLoS ONE**, v. 7, n. 11, p. 1–10, 2012.

FONSECA, B. D. et al. La-related protein 1 (LARP1) represses terminal oligopyrimidine (TOP) mRNA translation downstream of mTOR complex 1 (mTORC1). **Journal of Biological Chemistry**, v. 290, n. 26, p. 15996–16020, 2015.

FRANCO, J. R. et al. Epidemiology of human African trypanosomiasis. **Clinical Epidemiology**, v. 6, n. 1, p. 257–275, 2014.

FREIRE, E. et al. The Role of Cytoplasmic mRNA Cap-Binding Protein Complexes in Trypanosoma brucei and Other Trypanosomatids. **Pathogens**, v. 6, n. 4, p. 55, 2017.

FREIRE, E. R. et al. The four trypanosomatid eIF4E homologues fall into two separate groups, with distinct features in primary sequence and biological properties. **Molecular and Biochemical Parasitology**, v. 176, n. 1, p. 25–36, 2011.

FREIRE, E. R. et al. eIF4F-like complexes formed by cap-binding homolog TbEIF4E5 with TbEIF4G1 or TbEIF4G2 are implicated in post-transcriptional regulation in Trypanosoma brucei. **RNA**, v. 20, p. 1272–1286, 2014.

GALLIE, D. R. The role of the poly (A) binding protein in the assembly of the cap-binding complex during translation initiation in plants. **Translation**, v. 0731, p. 1–14,

2014.

GAUDENZI, J. DE; FRASCH, A. C.; CLAYTON, C. RNA-Binding Domain Proteins in Kinetoplastids : a Comparative Analysis. **Eukaryotic cell**, v. 4, n. 12, p. 2106–2114, 2005.

GEBAUER, F.; HENTZE, M. W. Molecular mechanisms of translational control. **Nat Rev Mol Cell Biol**, v. 5, n. 10, p. 827–35, 2004.

GIORDANI, F. et al. The animal trypanosomiases and their chemotherapy: a review. **Parasitology**, v. 143, n. 14, p. 1862–1889, 2016.

GLISOVIC, T. et al. {RNA}-binding proteins and post-transcriptional gene regulation. **FEBS Lett**, v. 582, n. 14, p. 1977–1986, 2008.

GOSS, D. J.; KLEIMAN, F. E. Poly(A) binding proteins: Are they all created equal? **Wiley Interdisciplinary Reviews: RNA**, v. 4, n. 2, p. 167–179, 2013.

GUERRA, N. et al. Identification and functional characterization of a poly(A)-binding protein from Leishmania infantum (LiPABP). **FEBS Letters**, v. 585, n. 1, p. 193–198, 2011.

GUERRERO, S. et al. Hiv-1 replication and the cellular eukaryotic translation apparatus. **Viruses**, v. 7, n. 1, p. 199–218, 2015.

HAILE, S.; PAPADOPOULOU, B. Developmental regulation of gene expression in trypanosomatid parasitic protozoa. **Current Opinion in Microbiology**, v. 10, n. 6, p. 569–577, 2007.

HAMILTON, T. L. et al. TOPs and their regulation. **Biochemical Society Transactions**, v. 34, n. 1, p. 12, 2006.

HARHAY, M. O. et al. Urban parasitology: Visceral leishmaniasis in Brazil. **Trends in Parasitology**, v. 27, n. 9, p. 403–409, 2011.

HASAN, A. et al. Systematic Analysis of the Role of RNA-Binding Proteins in the Regulation of RNA Stability. **PLoS Genetics**, v. 10, n. 11, p. 1–15, 2014.

HINNEBUSCH, A. G. The Scanning Mechanism of Eukaryotic Translation Initiation. **Annual Review of Biochemistry**, v. 83, n. 1, p. 779–812, 2014.

IMATAKA, H.; GRADI, A.; SONENBERG, N. A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. **EMBO Journal**, v. 17, n. 24, p. 7480–7489, 1998.

JACKSON, A. P. The evolution of amastin surface glycoproteins in trypanosomatid parasites. **Molecular Biology and Evolution**, v. 27, n. 1, p. 33–45, 2010.

JENSEN, B. C. et al. Extensive stage-regulation of translation revealed by ribosome profiling of Trypanosoma brucei. **BMC Genomics**, v. 15, n. 1, p. 911, 2014.

KAYE, P.; SCOTT, P. Leishmaniasis: complexity at the host-pathogen interface. **Nature reviews. Microbiology**, v. 9, n. 8, p. 604–615, 2011.

KAZEMI, B. Genomic organization of Leishmania species. **Iranian Journal of Parasitology**, v. 6, n. 3, p. 1–18, 2011.

KHACHO, M. et al. eEF1A Is a Novel Component of the Mammalian Nuclear Protein Export Machinery. **Molecular Biology of the Cell**, v. 19, n. 12, p. 5296–5308, 2008.

KHANAM, T. et al. Poly(A)-binding protein binds to A-rich sequences via RNA-binding domains 1+2 and 3+4. **RNA biology**, v. 3, n. 4, p. 170–177, 2006.

KOLEV, N. G.; ULLU, E.; TSCHUDI, C. The emerging role of RNA-binding proteins in the life cycle of Trypanosoma brucei. **Cellular Microbiology**, v. 16, n. 4, p. 482–489, 2014.

KOZLOV, G. et al. Structure and function of the C-terminal PABC domain of human poly(A)-binding protein. **Proceedings of the National Academy of Sciences of the United States of America**, v. 98, n. 8, p. 4409–4413, 2001.

KOZLOV, G.; GEHRING, K. Molecular basis of eRF3 recognition by the MLLE domain of poly(A)-binding protein. **PLoS ONE**, v. 5, n. 4, p. 3–8, 2010.

KRAMER, S. et al. Differential localization of the two T. brucei poly(A) binding proteins to the nucleus and RNP granules suggests binding to distinct mRNA pools. **PLoS.One.**, v. 8, n. 1932- 6203 (Electronic), p. e54004, 2013.

KUERSTEN, S.; GOODWIN, E. B. The power of the 3’ UTR: translational control and development. **Nat Rev Genet**, v. 4, n. 8, p. 626–637, 2003.

KUHN, U. et al. Xenopus poly(A) binding protein: functional domains in RNA binding and protein-protein interaction. **J.Mol.Biol.**, v. 256, p. 20–30, fev. 1996.

KÜHN, U.; WAHLE, E. Structure and function of poly(A) binding proteins. **Biochimica et Biophysica Acta - Gene Structure and Expression**, v. 1678, n. 2–3, p. 67–84, 2004.

LEDDA, M. et al. Effect of 3’ UTR length on the translational regulation of 5’-terminal oligopyrimidine mRNAs. **Gene**, v. 344, p. 213–220, 2005.

LIANG, X. et al. trans and cis Splicing in Trypanosomatids : Mechanism , Factors , and Regulation. **Eukaryotic cell**, v. 2, n. 5, p. 830–840, 2003.

LIN, J. et al. Nanopore detachment kinetics of poly(A) binding proteins from RNA molecules reveals the critical role of C-terminus interactions. **Biophysical Journal**, v. 102, n. 6, p. 1427–1434, 2012.

LOERCH, S.; KIELKOPF, C. L. Dividing and Conquering the RNA Recognition Motif Family: A Representative Case Based on hnRNP L. **J Mol Biol.**, v. 427, n. 19, p.

2997–3000, 2016.

LOPES, A. H. et al. Trypanosomatids: Odd Organisms, Devastating Diseases. **The Open Parasitology Journal**, v. 4, n. 1, p. 30–59, 2010.

LUEONG, S. et al. Gene expression regulatory networks in Trypanosoma brucei: Insights into the role of the mRNA-binding proteome. **Molecular Microbiology**, v. 100, n. 3, p. 457–471, 2016.

LUI, K. RanGTPase: A Key Regulator of Nucleocytoplasmic Trafficking. **Molecular and Cellular Pharmacology**, v. 1, n. 3, p. 148–156, 2009.

MARONDEDZE, C. et al. The RNA-binding protein repertoire of Arabidopsis thaliana. **Scientific Reports**, v. 6, p. 1–13, 2016.

MARTÍNEZ-CALVILLO, S. et al. Gene expression in trypanosomatid parasites. **Journal of Biomedicine and Biotechnology**, v. 2010, p. 525241, 2010.

MATTHEWS, K. R. The developmental cell biology of Trypanosoma brucei. **Cell**, v. 118, n. Pt 2, p. 283–290, 2009.

MCCONVILLE, M. J. et al. Living in a phagolysosome; metabolism of Leishmania amastigotes. **Trends in Parasitology**, v. 23, n. 8, p. 368–375, 2007.

MELO, E. O. et al. Identification of a C-terminal poly(A)-binding protein (PABP)-PABP interaction domain: Role in cooperative binding to poly(A) and efficient cap distal translational repression. **Journal of Biological Chemistry**, v. 278, n. 47, p. 46357–46368, 2003.

MERRICK, W. C. eIF4F: A retrospective. **Journal of Biological Chemistry**, v. 290, n. 40, p. 24091–24099, 2015.

MICHAELI, S. Trans-splicing in trypanosomes: machinery and its impact on the parasite transcriptome. **Future microbiology**, v. 6, n. 4, p. 459–474, 2011.

MONGE-MAILLO, B.; LÓPEZ-VÉLEZ, R. Treatment of Visceral Leishmaniasis. In: **The Leishmaniases: Old Neglected Tropical Diseases**. p. 169–190, 2018.

MOURA, D. M. N. et al. Two related trypanosomatid eIF4G homologues have functional differences compatible with distinct roles during translation initiation. **RNA biology**, 2015.

MÜLLER-MCNICOLL, M.; NEUGEBAUER, K. M. How cells get the message: dynamic assembly and function of mRNA-protein complexes. **Nature reviews. Genetics**, v. 14, n. 4, p. 275–287, 2013.

MYLER, P. J. et al. Leishmania major Friendlin chromosome 1 has an unusual distribution of protein-coding genes. **Proceedings of the National Academy of Sciences**, v. 96, p. 1–5, 1999.

NAJAFABADI, H. S. et al. Global identification of conserved post-transcriptional regulatory programs in trypanosomatids. **Nucleic Acids Research**, v. 41, n. 18, p. 8591–8600, 2013.

NILSSON, D. et al. Spliced leader trapping reveals widespread alternative splicing patterns in the highly dynamic transcriptome of Trypanosoma brucei. **PLoS Pathogens**, v. 6, n. 8, p. 21–22, 2010.

PHILIPPE, L. et al. La-related protein 1 (LARP1) repression of TOP mRNA translation is mediated through its cap-binding domain and controlled by an adjacent regulatory region. **Nucleic Acids Research**, v. 46, n. 3, p. 1457–1469, 2018.

PITULA, J.; RUYECHAN, W. T.; WILLIAMS, N. Trypanosoma brucei : Identification and Purification of a Poly (A)-Binding Protein. **Experimental Parasitology**, v. 160, n. 88, p. 157–160, 1998.

POMERANZ-KRUMMEL, D.; NAGAI, K. RNA-Binding Domains in Proteins. In: **Encyclopedia of Genetics**. p. 1733–1735, 2001.

PONTE-SUCRE, A. An overview of trypanosoma brucei infections: An intense host-parasite interaction. **Frontiers in Microbiology**, v. 7, n. DEC, p. 1–12, 2016.

PRÉVÔT, D.; DARLIX, J. L.; OHLMANN, T. Conducting the initiation of protein synthesis: The role of eIF4G. **Biology of the Cell**, v. 95, n. 3–4, p. 141–156, 16 maio 2003.

REQUENA, J. M. Lights and shadows on gene organization and regulation of gene expression in Leishmania. **Front Biosci**, v. 16, p. 2069–2085, 2011.

ROMANIUK, M. A.; CERVINI, G.; CASSOLA, A. Regulation of RNA binding proteins in trypanosomatid protozoan parasites. **World journal of biological chemistry**, v. 7, n. 1, p. 146–57, 2016.

SAKAMOTO, K. M. et al. Protacs: Chimeric molecules that target proteins to the Skp1-Cullin-F box complex for ubiquitination and degradation. **Proceedings of the National Academy of Sciences**, v. 98, n. 15, p. 8554–8559, 2001.

SHAHBABIAN, K. et al. Co-transcriptional recruitment of Puf6 by She2 couples translational repression to mRNA localization. **Nucleic Acids Research**, v. 42, n. 13, p. 8692–8704, 2014.

SILVA, J. R. DA C. **Análise da expressão heteróloga de proteínas com domínios de ligação a RNA em Leishmania infantum**. Fundação Oswaldo Cruz, 2014.

SLADIC, R. T. et al. Human PABP binds AU-rich RNA via RNA-binding domains 3 and 4. **Eur.J.Biochem.**, v. 271, p. 450–457, 2004.

SMITH, T. K. et al. Metabolic reprogramming during the Trypanosoma brucei life cycle. **F1000Research**, v. 6, p. 1–12, 2017.

SNUSTAD, D. P.; SIMMONS, M. J. **Fundamentos de Genética**. 6ª edição ed. Rio de Janeiro: Guanabara Koogan, 2013.

STUART, K. et al. Kinetoplastids: Related protozoan pathogens, different diseases. **Journal of Clinical Investigation**, v. 118, n. 4, p. 1301–1310, 2008.

TAKEUCHI, O.; YAMASHITA, A. Translational control of mRNAs by 3'-Untranslated region binding proteins. **BMB Reports**, v. 50, n. 4, p. 194–200, 2017.

TEIXEIRA, S. M. et al. Trypanosomatid comparative genomics: Contributions to the study of parasite biology and different parasitic diseases. **Genetics and Molecular Biology**, v. 35, n. 1, p. 1–17, 2012.

ULIANA, S. R. B.; RUIZ, J. C.; CRUZ, A. K. Leishmania Genomics : Where Do We Stand? In: **Bioinformatics in Tropical Disease Research: A Practical and Case-Study Approach**, p. 235-252, 2008.

WHEELER, R. J.; GLUENZ, E.; GULL, K. The limits on trypanosomatid morphological diversity. **PLoS ONE**, v. 8, n. 11, 2013.

WHO. **World Health Organization - Leishmaniasis**. Disponível em: <https://www.who.int/leishmaniasis/en/>.

YOFFE, Y. et al. Cap-binding activity of an eIF4E homolog from Leishmania. **RNA (New York, N.Y.)**, v. 10, n. 11, p. 1764–1775, 2004.

YOFFE, Y. et al. Binding specificities and potential roles of isoforms of eukaryotic initiation factor 4E in Leishmania. **Eukaryotic Cell**, v. 5, n. 12, p. 1969–1979, 2006.

YOFFE, Y. et al. Evolutionary changes in the Leishmania eIF4F complex involve variations in the eIF4E-eIF4G interactions. **Nucleic Acids Research**, v. 37, n. 10, p. 3243–3253, 2009.

ZOLTNER, M. et al. Comparative proteomics of the two T. brucei PABPs suggests that PABP2 controls bulk mRNA. **PLoS Neglected Tropical Diseases**, v. 12, n. 7, p. 1–18, 2018.

APÊNDICE A – ARTIGO PUBLICADO I

Fosforilação e interações associadas com o controle da função da proteína de ligação à cauda poli-A (PABP1) de *Leishmania* durante a iniciação da tradução

Artigo anexo na página seguinte.

A doutoranda e co-autora desse artigo publicado na *RNA biology*, foi responsável pela produção de extratos citoplasmáticos e imunoprecipitação da PABP1 e seus mutantes. Resumidamente, células de *Leishmania infantum* selvagem (controle negativo), com expressão heteróloga da PABP1 fusionada a epítopo HA e de mutantes dessa proteína, foram lisadas por cavitação para obtenção de extratos citoplasmáticos. Esses foram usados em dois experimentos independentes de imunoprecipitação. As proteínas parceiras co-imunoprecipitadas com a PABP1 e com seus mutantes (dados não inclusos no artigo) foram identificados por espectrometria de massas. Apenas aqueles polipeptídios enriquecidos no mínimo oito vezes, em relação ao controle negativo, foram considerados. Seis polipeptídios foram co-purificados com a PABP1: a própria PABP1; uma proteína de ligação ao RNA, a RBP23; os fatores de iniciação da tradução EIF4G3 e EIF4E4; e mais duas proteínas não caracterizadas, uma com os domínios NFT2 e RRM, e a outra contendo os domínios *Zinc Finger* e KH, a ZC3H41.

RNA BIOLOGY, 2018
VOL. 0, NO. 0, 1–17
https://doi.org/10.1080/15476286.2018.1445958

Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER

Phosphorylation and interactions associated with the control of the Leishmania Poly-A Binding Protein 1 (PABP1) function during translation initiation

Osvaldo P. de Melo Neto[a], Tamara D. C. da Costa Lima[b], Kleison C. Merlo[a], Tatiany P. Rom~ao[a], Pollyanna O. Rocha[a], Ludmila A. Assis[a], Larissa M. Nascimento[a], Camila C. Xavier[a], Antonio M. Rezende[a], Christian R. S. Reis[a] and Barbara Papadopoulou[c]

[a]Instituto Aggeu Magalh~aes – FIOCRUZ, Recife, PE, Brazil; [b]Centro Universitario Tabosa de Almeida, UNITA, Caruaru, Pernambuco, Brazil; [c]CHU de Quebec Research Center and Department of Microbiology-Infectious Disease and Immunology, Laval University, Quebec, QC, Canada

ABSTRACT

The Poly-A Binding Protein (PABP) is a conserved eukaryotic polypeptide involved in many aspects of mRNA metabolism. During translation initiation, PABP interacts with the translation initiation complex eIF4F and enhances the translation of polyadenylated mRNAs. Schematically, most PABPs can be divided into an N-terminal RNA-binding region, a non-conserved linker segment and the C-terminal MLLE domain. In pathogenic Leishmania protozoans, three PABP homologues have been identified, with the first one (PABP1) targeted by phosphorylation and shown to co-immunoprecipitate with an eIF4F-like complex (EIF4E4/EIF4G3) implicated in translation initiation. Here, PABP1 phosphorylation was shown to be linked to logarithmic cell growth, reminiscent of EIF4E4 phosphorylation, and coincides with polysomal association. Phosphorylation targets multiple serine-proline (SP) or threonine-proline (TP) residues within the PABP1 linker region. This is an essential protein, but phosphorylation is not needed for its association with polysomes or cell viability. Mutations which do impair PABP1 polysomal association and are required for viability do not prevent phosphorylation, although further mutations lead to a presumed inactive protein largely lacking phosphorylated isoforms. Co-immunoprecipitation experiments were carried out to investigate PABP1 function further, identifying several novel protein partners and the EIF4E4/EIF4G3 complex, but no other eIF4F-like complex or subunit. A novel, direct interaction between PABP1 and EIF4E4 was also investigated and found to be mediated by the PABP1 MLLE binding to PABP Interacting Motifs (PAM2) within the EIF4E4 N-terminus. The results shown here are consistent with phosphorylation of PABP1 being part of a novel pathway controlling its function and possibly translation in Leishmania.

ARTICLE HISTORY
Received 31 July 2017
Revised 21 December 2017
Accepted 13 February 2018

KEYWORDS
Leishmania; translation; phosphorylation; protein-protein interaction; translation initiation factor; polyadenylate-binding protein

Introduction

The trypanosomatids constitute a group of parasitic flagellated protozoans which include major human pathogens belonging to the genera Leishmania and Trypanosoma. These organisms are characterized by a complex biology and singular molecular aspects rarely encountered in other eukaryotes. Unique features include polycistronic transcription and regulation of gene expression based mainly on post-transcriptional mechanisms, targeting events such as mRNA stability and translation [1-7]. A major target for regulation is possibly the initiation stage of protein synthesis where a number of eukaryotic translation initiation factors are known to act [8]. Paramount among those is the translation initiation factor eIF4F, a complex of three distinct subunits including the RNA helicase eIF4A, the cap binding protein eIF4E and the scaffolding subunit eIF4G. During translation, eIF4F binds to the mRNA and facilitates the recruitment of the small ribosomal subunit to its 5^0 end [9-12]. A key eIF4F partner is the cytoplasmic Poly-A Binding Protein (PABPC or simply PABP), a conserved eukaryotic protein involved in multiple cellular functions associated with the metabolism of messenger RNAs, including mRNA polyadenylation/deadenylation, export, surveillance, degradation, and translation [13-20]. At the initiation stage of protein synthesis, PABP is responsible for the enhanced translation of polyadenylated mRNAs, facilitating their recruitment by the translation machinery. In model eukaryotes, this is accomplished through a direct interaction between PABP and eIF4F. In mammals, yeasts and plants the interaction between eIF4F and PABP is mediated through the eIF4G subunit [10,13,14,16,20].

Schematically, PABPs are formed by a N-terminal region, a non-conserved linker segment and the unique C-terminal MLLE domain, originally called PABC (for reviews see [10,13,14,16,20]). The N-terminal region consists of roughly two-thirds of the pro-tein and contains four conserved RNA binding domains (RRMs) positioned in tandem, each defined by the presence of two highly conserved RNP motifs (RNP1 and RNP2) [21,22]. The PABP RRMs 1 and 2 are responsible for the specific binding to poly-A and also mediate the protein's interaction with eIF4G [23-26].

CONTACT Osvaldo P. de Melo Neto opmn@cpqam.fiocruz.br Departamento de Microbiologia, Instituto Aggeu Magalh~aes/Fiocruz. Av. Prof. Moraes Rego, s/n – Cidade Universitaria – Campus da UFPE, Recife, PE, 50740–465, Brazil; Barbara Papadopoulou barbara.papadopoulou@crchul.ulaval.ca Department of Microbiology and Immunology, Laval University Research Center in Infectious Disease, RC-709, CHUL Research Center (CHUQ) 2705 Laurier Blvd. Quebec, QC, G1V 4G2, Canada.

Supplemental data for this article can be accessed on the publisher's website.

RRMs 3 and 4 seem to bind to non-polymeric AU sequences with unknown functions and also mediate protein-protein interactions [27-29]. In mammals, the PABP linker region seems to be involved in the multimerization process where multiple PABP molecules associate with each other attached to the poly-A tail [30,31]. As for the C-terminal MLLE domain, it mediates the interaction of PABP with other polypeptides having a conserved peptide motif named PAM2 [32-34].

In Leishmania three PABP homologues were identified but only two of those (PABP1 and PABP2) are found conserved in most, if not all, trypanosomatids [35,36]. It has been shown that all three Leishmania PABPs and T. brucei PABP2 have affinity to poly-A, although the PABP2 orthologues share sub-stitutions in conserved amino acids residues that are critical for poly-A recognition [35,37,38]. All three Leishmania PABPs are abundant proteins which are constitutively expressed and local-ize to the cytoplasm although PABP2 and PABP3, but not PABP1, migrate to the nucleus upon inhibition of transcription. Co-precipitation studies have also shown a strong association between PABP2 and PABP3 which likely bind to the same set of target mRNAs [35]. In T. brucei, through tethering assays, both PABP1 and PABP2 have been shown to stimulate transla-tion of a reporter mRNA when tethered to its 3 ^{0} end [39,40] and have been found to be associated with polysomes [36,41]. RNAi depletion of the T. brucei PABP1 and PABP2 ortho-logues confirmed that both proteins are essential for cell viabil-ity [35], however the two proteins localize to distinct sets of inducible RNP granules and PABP2, but not PABP1, can accu-mulate in the nucleus upon inhibition of mRNA maturation and heat shock treatment [36].

Multiple homologues have also been identified in trypanoso-matids for the eIF4E and eIF4G subunits of eIF4F. The six con-served eIF4Es (EIF4E1 through EIF4E6) and five conserved eIF4Gs (EIF4G1 through EIF4G5) differ substantially in sequence and in binding partners [42-46]. Two distinct eIF4F-like complexes, centered on the interactions between EIF4E4/ EIF4G3 and EIF4E3/EIF4G4, have been characterized with properties which implicate them during translation initiation [47-51]. In both Leishmania and Trypanosoma species, the EIF4E4/EIF4G3 complex has been shown to be associated with PABP1 [35,36,49] and the Leishmania PABP1 has been seen to bind directly to EIF4E4, through an interaction unknown from other eukaryotes [49]. The novel PABP1/ EIF4E4 interaction seems to be stronger than the one between EIF4E4 and EIF4G3 and it was also shown to be more critical for EIF4E4 function in vivo [52].

The Leishmania PABP1 was early on seen to be targeted by phosphorylation events, resulting in a clearly identifiable phos-phorylated isoform which was seen to be reduced upon tran-scription inhibition [53]. In contrast, no similar isoforms have been detected for PABP2 or PABP3 [35], implying the existence of specific regulation directed to PABP1. In contrast, EIF4E4, the PABP1 partner, has also been found to be phosphorylated [52], highlighting the two proteins as targets for regulatory mechanisms controlling their functions. Here, we have investi-gated PABP1 in more detail, assessing its phosphorylation as well as interactions with EIF4E4 and other protein partners, with the purpose of clarifying aspects of its function in Leish-mania and other trypanosomatids and enhancing our knowledge on translation initiation in these organisms. Using wild-type and mutated PABP1 variants in L. infantum, pro-tein-protein interaction assays and complementation studies after gene knockout, we defined a clear pattern of phosphoryla-tion for PABP1, which is strikingly similar to the one previ-ously reported for EIF4E4 [52]. We further investigated requirements for phosphorylation as well as its role for PABP1 function and cell viability. Binding partners were identified as well as residues required for the protein to function properly. The novel EIF4E4/PABP1 interaction was further characterized in detail and shown to involve PAM2 like motifs in EIF4E4 and the MLLE domain in PABP1.

Results

PABP1 is constitutively expressed across the L. infantum life cycle but is hyperphosphorylated during growth conditions of active translation

PABP1 orthologues from both T. brucei and Leishmania species have been shown to be targeted by phosphorylation events lead-ing to multiple isoforms which, nevertheless, were not clearly associated with any specific growth stage or biological process [35,36,53]. In T. brucei, in a high throughput search for phos-phoproteins, both PABP1 and PABP2 were found to be phos-phorylated, with multiple residues targeted for phosphorylation in each protein [54] but only PABP1 was represented by more than one isoform upon expression analysis using whole parasite lysates [36]. Here, we first compared the expression of Leish-mania PABP1 and its phosphorylated isoforms with PABP2 dur-ing standard promastigote growth conditions. In Leishmania, the original experiments focusing on the characterization of PABP1 were carried out in L. major [35] but here we opted to use a sec-ond Leishmania species, L. infantum, which is more amenable to differentiation in culture. Growth curves were set up starting with stationary-phase promastigotes passaged to new media and evaluated for the expression of the two Leishmania PABP homo-logues, at selected time points, using rabbit polyclonal antiserum directed against the two proteins. As shown in Fig. 1A, both pro-teins are constitutively expressed throughout all stages of the growth curve assayed, but only PABP1 is clearly represented by more than one isoform. In contrast to the constant pattern of expression seen for PABP2, the expression of the two detected PABP1 isoforms vary during the growth curve, with the top (phosphorylated) band being absent or nearly absent from sta-tionary phase cells (days 5 to 6), but predominating soon after passaging (2 hours) and during the stages of fast, logarithmic, cell growth (up to 2 days).

Next, we investigated the PABP1 expression during amasti-gote differentiation using two consecutive growth curves (Fig. 1B). The differentiation curve started with stationary-phase promastigotes, which were allowed to differentiate in amastigote medium and it was followed by the second curve using fully dif-ferentiated cells passaged again into the new amastigote medium. As seen for promastigotes, the expression pattern of the PABP1 isoforms varied for both curves also according to the growth stage, with the top band being predominant soon after passaging and during logarithmic growth and the bottom band being more abundant in stationary cells. The changes in isoform profile for

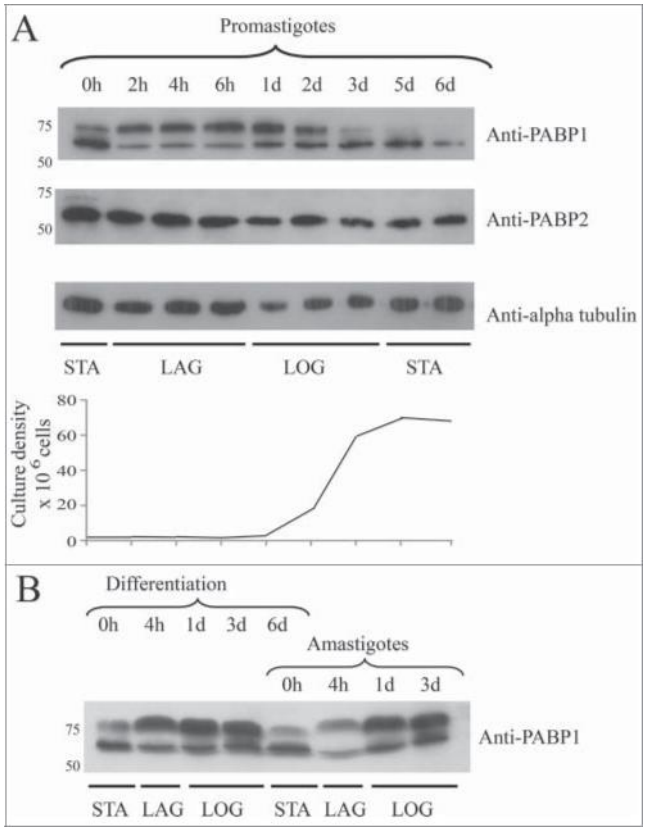


Figure 1. Expression analysis of L. infantum PABP1 during both promastigote and amastigote life stages. (A) Western blotting comparing the PABP1 expression with that of PABP2 and alpha-tubulin during distinct promastigote growth phases. The results from a single growth curve are shown with aliquots taken immediately after passaging to start a new culture (0h), at 2, 4 and 6 hour time points and daily after that. For all lanes equal loads were run under denaturing conditions and blotted with whole rabbit polyclonal sera directed against Leishmania PABP1 and PABP2 or serum against alpha tubulin (protein loading control). The results shown are representative of multiple experiments carried out not only with L. infantum, but also with L. major and L. amazonensis. The graphic representation of the cell counts from every aliquot is shown below the blot. (B) Western blotting showing the PABP1 expression in L. infantum during two consecutive curves grown in amasti-gote media. The 1st curve, the differentiation curve, started with stationary phase promastigotes and the second curve started with fully differentiated, stationary phase amastigotes as previously described [52] and, as previously shown, the dif-ferentiation into amastigotes was confirmed through the detection of the A2 amastigote-specific marker. STA – Stationary cells. LAG – Lag phase culture. LOG – Logarithmic phase culture. For the PABP blots relevant molecular weight markers are shown on the left.

PABP1 is then clearly linked with the growth phase of the cells and is not influenced by its differentiation stage. This pattern of isoform expression observed for PABP1 and common to both promastigote and amastigote stages of the Leishmania life cycle is very similar to or identical to the one seen for EIF4E4, PABP1's binding partner. As previously reported, EIF4E4 is targeted by multiple phosphorylation events generating multiple isoforms which predominate in both promastigote and amastigote forms during logarithmic growth and presumably active protein synthe-sis [52,55].

PABP1 phosphorylation is linked to SP/TP motifs found within its linker region

L. infantum PABP1 has been shown to be phosphorylated in vitro by a p38 MAP kinase [38]. The very fast changes in PABP1 isoform expression induced by the passaging to new media are also consistent with a MAP kinase response. In T. brucei the three PABP1 residues that were found to be targeted by phos-phorylation are serine or threonine residues followed by a pro-line [54], known targets of MAP kinases. To investigate possible phosphorylation sites within PABP1, and considering the simi-larity observed with the EIF4E4 phosphorylation targeting con-served serine-proline (SP) or threonine-proline (TP) motifs, we performed an alignment of multiple PABP1 orthologues from different trypanosomatid species and searched for conserved SP or TP motifs that are missing from equivalent regions from PABP2 and PABP3 (Fig. S1). Seven of these motifs were found in close proximity at the end of the linker region of PABP1 sequences from all Leishmania species investigated. These are indicated in the scheme from Fig. 2A and highlighted in the alignment, comparing the distantly related L. infantum and L. braziliensis PABPs. Three of the motifs are also conserved in T. brucei and they coincide with the previously identified phos-phorylation sites in PABP1 from this organism [54]. The align-ment also highlights the three motifs found within the T. brucei PABP1 (Fig. S1).

To investigate the identified motifs further, we opted here to use site directed mutagenesis in order to mutate, within the L. infantum PABP1 gene, the sequences encoding six out of the seven candidate SP/TP motifs found in the PABP1 linker region (circled in Fig. 2A). Each mutation resulted in the sub-stitution of the serine/threonine by an alanine, effectively pre-venting the phosphorylation of the motif. The mutant protein (named TP-SP Mut) was then ectopically expressed after stable transfection of the mutated gene into L. infantum, generating a recombinant protein tagged with the HA epitope at its C-termi-nus. First, we compared its expression with that of an equiva-lent HA-tagged wild-type PABP1 (from now own named HA-PABP1). After blotting with an anti-HA monoclonal antibody, the HA-PABP1protein displayed the same isoform expression pattern seen for the endogenous protein (Fig. 2B). A lower MW band was seen predominantly at the beginning of the growth curve and also at day 5 of the culture, when cells reach stationary phase. In contrast, during logarithmic growth at day 2, a higher MW band, corresponding to the phosphorylated PABP1 isoform, was detected (Fig. 2B). The HA-tagged TP-SP mutant was also expressed and assayed under the same growth conditions as the wild-type protein (Fig. 2B, TP-SP Mut). A single protein band was seen for the HA-tagged mutant, with no changes observed in size or expression between logarith mically grown or stationary cells, indicating a lack of phosphor-ylation throughout the growth curve. These results are consis-tent with multiple phosphorylation events targeting PABP1 and directed at the multiple conserved SP/TP motifs found in its linker region. These phosphorylation events lead to the high molecular weight isoform observed in the PABP1 expression analysis (Fig. 1) and are possibly mediated by the same kinase (s) which target EIF4E4.

PABP1 is enriched in polysomal fractions during Leishmania logarithmic growth

We have previously attempted to correlate the presence of the phosphorylated EIF4E4 isoforms with a more active role in

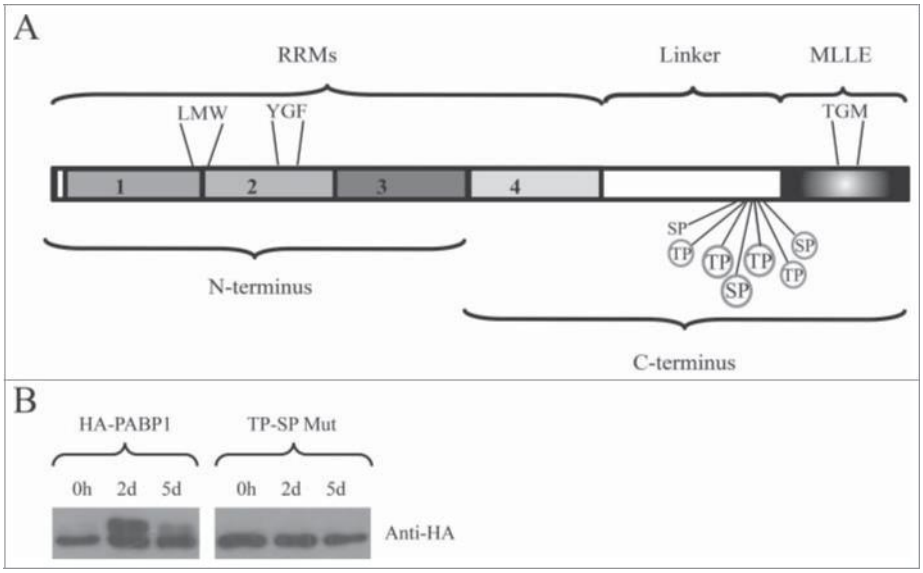


Figure 2. Identification of the PABP1 phosphorylation motifs. (A) Structural domain organization of the L. infantum PABP1 homologue. The protein's RRM region consist-ing of the first two thirds of PABP1 (encompassing all four RRMs) plus the "Linker" segment and the "MLLE" domain are shown. Seven putative phosphorylation sites are indicated, six of which were targeted by site-directed mutagenesis (circled). Also highlighted are the three sets of amino acid triplets chosen for mutagenesis and local-ized between RRMs 1 and 2 (LMW), within RRM2 (YGF) or within the MLLE domain (TGM) The N-terminus and C-terminus fragments used for the pull-down assays are also indicated. (B) Expression of HA-tagged, episomally encoded, PABP1 evaluated with a monoclonal commercial anti-HA antibody. The left panel compared the expres-sion of wild-type PABP1 (HA-PABP1) during three representative stages of the parasite growth curve. The right panel evaluates the expression, under the same conditions of the PABP1 mutant (TP-SP Mut) where all six putative phosphorylation sites were targeted by site directed mutagenesis.

translation by investigating its association with mRNA bound polysomes. Unfortunately, in our hands at least, the very labile nature of this protein and its propensity to degradation during preparation of these extracts prevented a clear identification of EIF4E4 after polysome analysis (unpublished data). PABP1, however, is more abundant than EIF4E4 and is not as susceptible to degradation, so its association with polysomes could be investigated. Sucrose gradient fractionation of cellular extracts from cells grown logarithmically and after reaching stationary phase was then performed. Extracts from logarithmically grown cells yielded profiles with multiple peaks for the polysomes, compatible with robust protein synthesis. For the stationary cells, in contrast, the polysome peaks found in the sucrose-gra-dient profiles were minimal or non-existent, an indication of very limited protein synthesis (Fig. S2).

Fractions from the top of the gradients ("Free mRNP") and from the 40S and 60S ribosomal subunits, 80S ribosomes and the polysomes were pooled and blotted with the PABP1 anti-serum. For the gradients derived from the logarithmically grown cells, PABP1 was found associated mostly with the 80S fraction and the polysomes (Fig. 3A). However, in gradients from stationary cells, PABP1 was found mostly on the top of the gradient with only a very limited amount found in the frac-tions corresponding to the polysomes, compatible with the low levels of protein synthesis at this stage. In these experiments the top PABP1 isoform (phosphorylated) was not clearly visi-ble, especially in the polysome fractions, and we believe this is likely due to dephosphorylation occurring during extract prep-aration and subsequent steps of gradient fractionation. Indeed when freshly prepared cytoplasmic extracts are compared with whole cell extracts, prepared by the resuspension of live cell directly in SDS-PAGE sample buffer, this isoform is already much reduced and this was seen even in the presence of phosphatase inhibitors (not shown). This contrasts to what has been described from T. brucei, where phosphorylated PABP1 can be detected both in polysomal and non-polysomal fractions [36]. In Fig. 3, a control blot was also carried out using a rabbit polyclonal antiserum directed against the ribosomal protein P0, a constituent of the 60S ribosomal subunit. As expected, P0 was

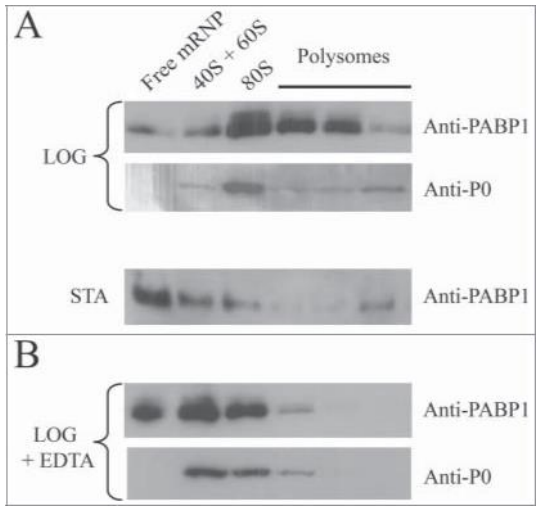


Figure 3. Polysome profile analysis of PABP1 in logarithmically growing and stationary cells. (A) Sucrose gradients (15-45%) for polysome profile analysis of L. infantum promastigotes were carried out using extracts from both logarithmic (LOG) or stationary (STA) phase stages of representative growth curves. Samples from 260 nm peaks derived from the free mRNPs, 40/60S and 80S ribosomal frac-tions, as well as the polysomes, were pooled and blotted with the anti-PABP1 anti-bodies and antibodies directed against the Leishmania ribosomal protein P0. (B) EDTA was used with the samples from logarithmic growth (LOG + EDTA) to dem-onstrate the dissociation of polysomes and corresponding effect on the migration of PABP1 and P0 on the sucrose gradients. The results shown are representative of a minimum of two different experiments carried out with different batches of cell growth.

detected on the polysome fractions from logarithmically grown cells as well as in the 40S/60S and 80S fractions (Fig. 3A). Another control was performed by treating extracts from loga-rithmically grown cells with EDTA prior to loading on the sucrose gradient, followed by blotting to assay for both PABP1 and P0 (Fig. 3B). Polysome dissociation was confirmed by the detection of P0 mainly in the 40S/60S and 80S fractions. PABP1, although more abundant, displayed a profile similar to P0 with most of it shifting to the top of the gradient upon EDTA treatment. Altogether, these results indicate that the Leishmania PABP1 protein is strongly associated with the poly-somes during logarithmic growth.

Identification of motifs relevant for PABP1 phosphorylation and polysome association

In order to better define the mechanism(s) associated with PABP1 phosphorylation and its role in translation, we opted to search for motifs that would be required for specific aspects of PABP1 function and that could be identified by mutagenesis. This also follows a similar approach carried out with EIF4E4 where, for instance, the elimination of its interaction with its EIF4G3 partner did not prevent its phosphorylation [52], con-trasting to mammalian systems where eIF4E phosphorylation requires binding to eIF4G, which in turn recruits the kinase Mnk [56]. Here, three sets of three consecutive amino acid resi-dues were chosen for mutagenesis and which are indicated in Fig. 2A. All three are strictly conserved in different PABP1 orthologues and are also mostly conserved between PABP homologues from different organisms, including metazoans, plants and yeast, but the first two are specifically modified in trypanosomatid PABP2 orthologues. The first triplet, leucine-methionine-tryptophan (LMW) maps to a segment between the PABP RRMs 1 and 2 which has been shown to mediate PABP binding to eIF4G and eIF4B homologues [10]. The sec-ond triplet, tyrosine-glycine-phenylalanine (YGF), is part of the RNP2 motif from RRM2 and includes residues which have been implicated in poly-A binding and recognition [22]. A third triplet, chosen within the conserved structural core of the MLLE domain [57], consists of threonine-glycine-methionine (TGM), and is the only one also found in PABP2. All mutations replaced the selected residues for alanines, generating the LMW/AAA, YGF/AAA and TGM/AAA mutants which were also expressed as HA-tagged proteins in transfected L. infan-tum cells. The expression of the respective recombinant pro-teins was then evaluated under conditions of logarithmic growth where the phosphorylated isoform of PABP1 predomi-nates (Fig. 4A). All three mutants were represented by the two isoforms seen with wild-type PABP1, with the top phosphory-lated isoform being in general predominant and no indication of an effect induced by the mutations on phosphorylation (Fig. 4A). A fourth mutant, containing the first two triplets mutated (LMW/AAA + YGF/AAA) was then generated and evaluated likewise. This double mutant seemed to be expressed at lower levels than the previous proteins, although no proper quantitation was carried out to evaluate this, and very little expression of the top isoform was seen, indicating a significant impact on PABP1 phosphorylation when both LMW and YGF motifs were mutated.

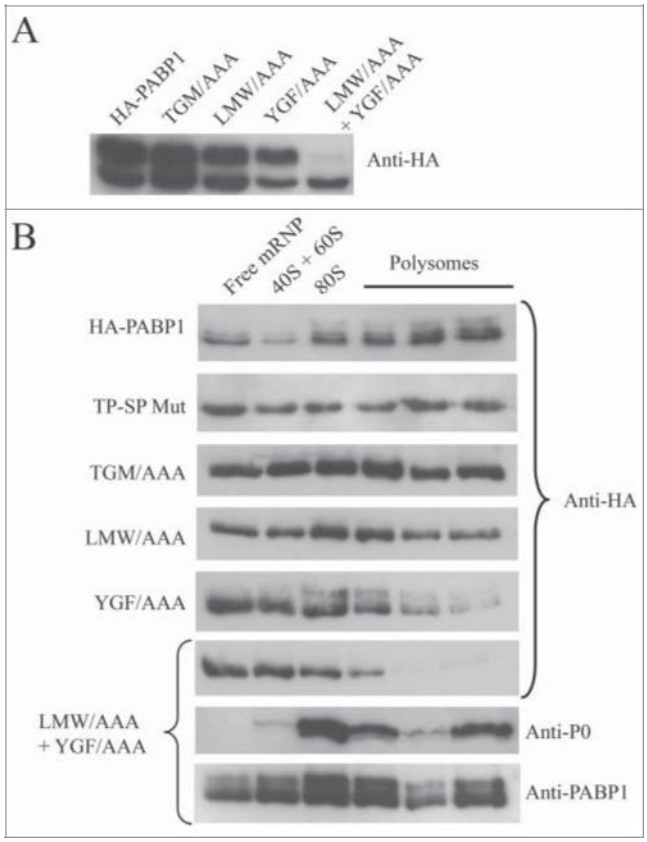


Figure 4. Identification of relevant motifs involved in PABP1 phosphorylation and polysome binding. (A) Expression analysis of HA-tagged PABP1 mutants containing selected triplets of amino acid residues replaced by alanines. The expression of the resulting proteins was evaluated in logarithmically grown cells, as shown at the 2d time point in Fig. 2B. The effect on PABP1 phosphoryla-tion of mutating the triplets LMW, YGF or TGM, individually, or both sets of LMW and YGF was evaluated. The results shown were reproduced using a minimum of two growth curves for each mutant and, as much as possible, the cultures were grown in parallel and assayed at the same stage of the growth, using the same number of passages following transfection. (B) Polysome profile analysis of L. infantum promastigotes expressing wild-type or mutant HA-tagged PABP1 pro-teins, under logarithmic growth, were carried out to investigate the association of each individual mutant to polysomes. Samples were pooled as shown in Fig. 3 and blotted with the anti-HA antibody. For the LMW/AAA + YGF/AAA mutant, as control, the same samples were also blotted using the anti-P0 and anti-PABP1 antisera.

To investigate if PABP1 phosphorylation was a requirement for it to bind to polysomes, wild-type and mutant HA-tagged proteins were evaluated through polysome profiling using sucrose gradients carried out as described in Fig. 3 with pro-mastigote cells grown logarithmically. As expected, wild-type HA-tagged PABP1 associated efficiently with the polysomes, with most of the ectopically expressed protein seen in the poly-somal fractions (Fig. 4B). A similar profile was seen for the TP-SP mutant, where phosphorylation was abolished, indicating that phosphorylation of PABP1 is not a requirement for its association with polysomes. The various mutants assayed in Fig. 4A were also investigated using the same approach. Both LMW/AAA and TGM/AAA mutants were also found with the polysomes although, when compared to the wild-type protein, these seemed to be found in greater quantities in the lighter non-polysomal fractions of the gradients, indicating a possible reduction in efficiency in polysome association (Fig. 4B). The YGF/AAA mutant, in contrast, was clearly impaired on its

ability to associate with mRNA bound to polysomes, presum-ably a consequence of the mutations targeting residues required for mRNA binding. This effect was further enhanced in the LMW/AAA + YGF/AAA double mutant, which showed no indication of association with polysomes with most of the pro-tein found on the top fractions of the gradients (Fig. 4B). For this last mutant, the same sucrose gradient fractions were also blotted with anti-sera directed against the P0 ribosomal protein and the native PABP1 (this anti-sera recognizes both the HA-tagged mutant as well as the endogenous PABP1 present in the extracts). Both P0 and native PABP1 were mainly found with the polysome fractions, confirming the quality of the sucrose gradient fractionation and that the lack of association of the double mutant with the polysomes was due to a disruption of poly(A) tract binding. In comparison to the isoforms observed in the whole cell extracts from Fig. 4A, phosphorylated iso-forms in the gradient fractions from different HA tagged pro-teins are much diminished in intensity and, when present, are reduced in apparent molecular weight. As for the native PABP1 shown in Fig. 3 we believe this is due to dephosphorylation and their presence/absence from different fractions does not necessarily have any functional implications. Nonetheless, the sucrose gradient results with the mutant proteins are consistent with phosphorylation not being required for PABP1 to be recruited by the translation apparatus. These experiments also highlight the functional relevance of the LMW and YGF motifs within the PABP1 RRM region.

The Leishmania PABP1 gene is essential for parasite growth

In T. brucei, RNAi mediated knockdown of PABP1 prevented cell viability and growth in procyclic cells confirming it is essential for survival in these cells [35]. However, the existence of a third PABP homologue in Leishmania species (PABP3), lost from the Trypanosoma lineage [36], means that the results from T. brucei cannot readily be extrapolated to all trypanoso-matids. Here, attempts were made to generate haploid or dip-loid knockouts for the PABP1 gene and evaluate its effect on L. infantum survival. The gene replacement strategy used for generating the knockouts is the same applied previously to study EIF4E4 [52], (see also Fig. 5A). The first PABP1 copy gene replacement was done using the hygromycin resistance gene (HYG) expression cassette for the generation of the hap-loid PABP1/PABP1::HYG cell line (HYG KO). Transfected cells were recovered with no difficulties followed by cloning in order to have multiple lineages having the single knockout (SKO) of the PABP1 gene. Integration into the endogenous PABP1 locus was confirmed through a PCR based analysis using a primer annealing to the PABP1 gene intergenic region, external to the fragment used for the integration procedure, and an internal primer annealing to the hygromycin coding sequence (Fig. 5B). The presence of the remaining PABP1 allele was also confirmed using a similar PCR procedure, as illustrated in Fig. 5. This sec-ond wild type allele was then targeted with a cassette expressing the puromycin resistance gene (PUR KO). Multiple transfection events were carried out using independent clones generated with the haploid genotype. These transfections did not always yield viable transfected cells but sometimes cell lines were recovered which had the PUR gene integrated into the PABP1 locus (PABP1::PUR /PABP1::HYG), suggesting a double knock-out of the PABP1 genes (DKO). When investigated through PCR, however, an intact copy of the endogenous PABP1 gene was still detected (not shown). Thus, our conclusion is that the PABP1 gene is essential for cell viability and susceptible to gene duplication events, leading to the generation of a trisomic gene copy upon targeted gene replacement of the two PABP1 alleles.

To evaluate if both copies of the PABP1 gene could be targeted without inducing the gene duplication event, and to eliminate the possibility of the PABP1 gene being originally trisomic, an alternative complementation strategy, also previ-ously applied to the EIF4E4 gene [52], was tested. This strat-egy requires transfecting in the HYG SKO cell lines a plasmid encoding the PABP1 gene (conferring resistance to neomycin) prior to the transfection with the puromycin expression cas-sette (PUR KO). Again, multiples HYG SKO cell lines were used for the transfection with the plasmid encoded PABP1 gene, followed by neomycin selection and subsequent trans-fection using the PUR KO cassette and selection with puromy-cin. Using this strategy, the second PABP1 allele was successfully replaced by the puromycin cassette, as confirmed by the lack of PCR amplification of the endogenous PABP1 gene fragment in the presence of the PUR KO gene fragment. For these gene replacement experiments, in order to indepen-dently confirm the deletion of the two endogenous copies of the PABP1 gene, we also opted to use Southern-blot, as previ-ously applied for EIF4E4. DNA from the wild-type L. infan-tum as well as the single (SKO) and double (DKO) knockout cells lines was probed with the gene fragment corresponding to the PABP1 gene 5^{0} intergenic region, a fragment missing from the plasmid encoded gene. As shown in Fig. 5C, the band corresponding to the endogenous PABP1 gene is miss-ing from the DKO cell lines, confirming the deletion of both gene copies by the knockout procedures. Altogether, the results from Fig. 5B and 5C confirm the essential nature of the PABP1 gene, at least for the promastigote stage of the Leishmania cell cycle.

Complementation studies reveal motifs strictly required for PABP1 function

Next, attempts were made to recover double knockout (DKO) cells from the HYG SKO lines expressing the different recombi-nant proteins in order to see which of the mutants could com-plement the lack of endogenous PABP1. Successful recovery of cell lines in the presence of both hygromycin and puromycin selection was achieved for cells expressing all mutant proteins, although, as seen for HYG SKO derived lines (discussed below), the recovery was significantly delayed in comparison with cells expressing the wild-type PABP1 (not shown). DNA recovered from all putative DKO lines was assayed, by PCR only, for the presence of the endogenous PABP1 gene and representative results are shown in Fig. 6A. Successful depletion of the endog-enous gene was reproducibly achieved for the cells expressing the TP-SP phosphorylation mutant as well as the one having the TGM/AAA mutation. In contrast, for those cells comple-mented with the LMW/AAA and YGF/AAA mutants, the endogenous PABP1 gene was still present after hygromycin

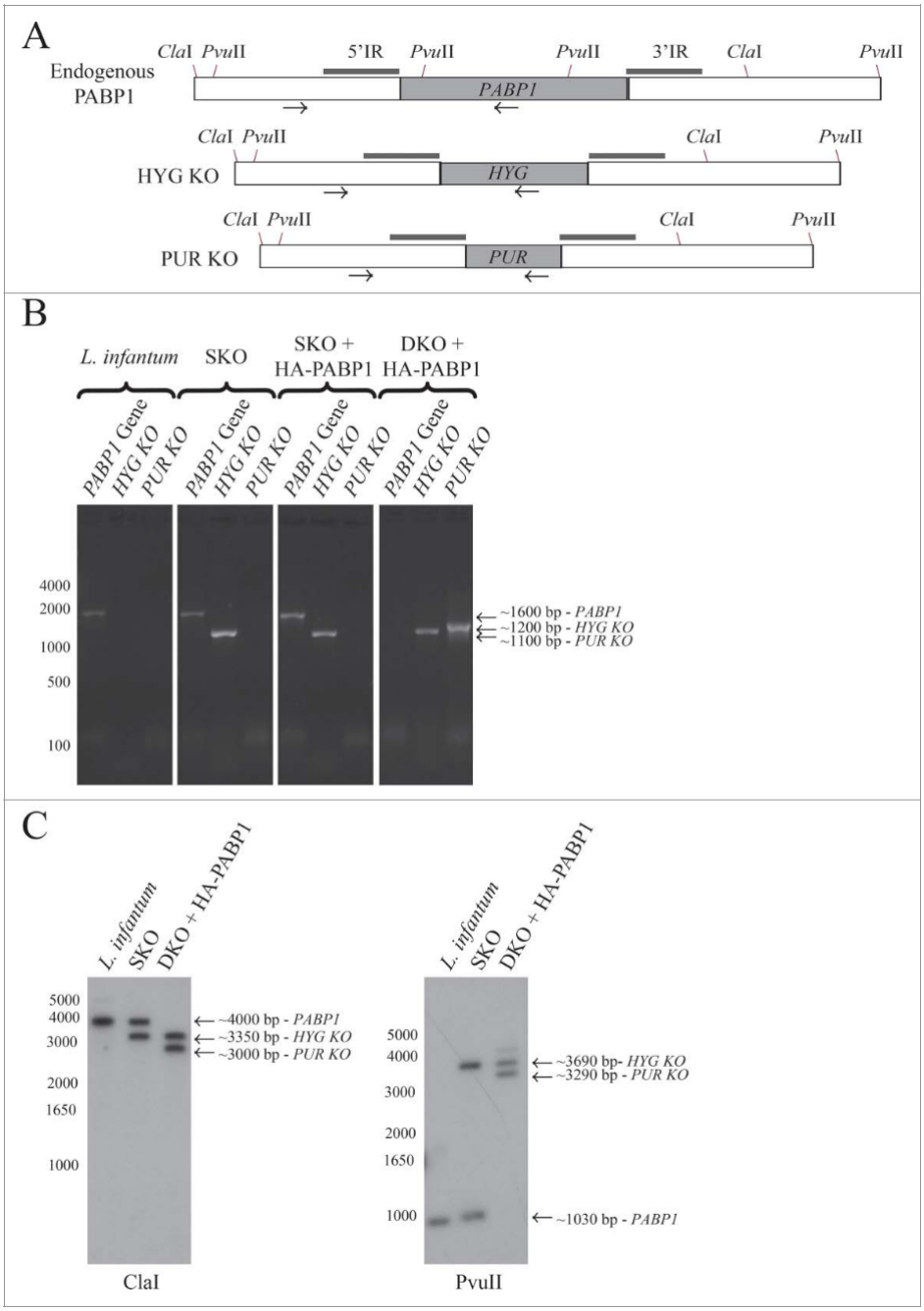


Figure 5. Generation of haploid and diploid PABP1 gene knockout mutants in L. infantum. (A) Schematic representation of the endogenous PABP1 gene as well as gene replacements following the single integration of the hygromycin (HYG) and/or puromycin (PUR) drug selectable markers. The 500 bp intergenic regions upstream and downstream of the PABP1 coding region (named 5 ^{0}IR and 3 ^{0} IR in the figure), and that were used to generate the knockout cell lines through homologous recombination, are indicated by the small rectangles in dark gray. The position of the primers added to the PCR reactions used to confirm the integration events are indicated by arrows, while the position of the restriction sites used for the Southern blots are also shown. (B) Agarose gel showing the PCR fragments amplified from the wild type L. infantum as well the single knockout (SKO) lineages and the double knockout (DKO) lineages generated in the presence of a plasmid encoded PABP1 gene. The oligonucleotides used for the PCR reactions are listed in Table S5. (C) Southern blots confirming the efficiency of the knockout procedures. Total genomic DNA from wild-type L. infantum as well as the SKO and DKO cells lines was digested with either ClaI or PvuII restriction enzymes and probed with the gene fragment corresponding to the PABP1 gene 5 ^{0} intergenic region (5 ^{0} IR). For (B) and (C), fragments corresponding to the wild-type PABP1 gene as well as the hygromycin (HYG KO) and puromycin (PUR KO) integration events are shown. Size markers in bp are shown on the left of the panels, while the sizes of the bands of interest are also indicated on the right. All transfection experi-ments were performed at least twice, generally more, with the DKOs performed using a minimum of two sets of SKO cell lines. Selected samples were used for the PCR or Southern blots.

and puromycin dependent integration events, indicating a gene duplication event similar to the one observed during our attempts to generate the PABP1::PUR /PABP1::HYG cell lines.

The absence of the endogenous PABP1 in the DKO lines complemented with the phosphorylation mutant was also assessed using the anti-PABP1 serum. As shown in the western blot of Fig. 6B, only the non-phosphorylated PABP1 isoform was detected in cells complemented with the TP-SP mutant in contrast to the cell line overexpressing the HA-tagged wild-type PABP1 which produces both phosphorylated and non-phosphorylated isoforms. Thus, a conclusion from the complementation experiments is that phosphorylation of

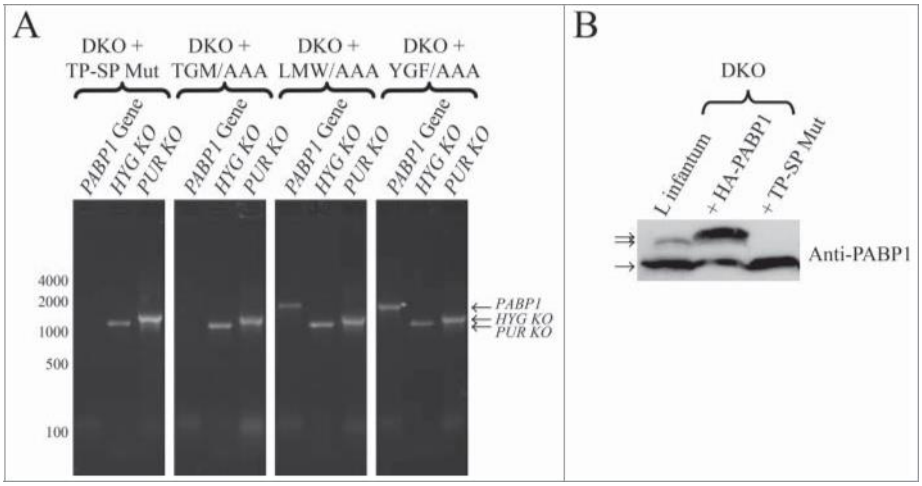


Figure 6. Evaluation of the diploid PABP1 knockout in L. infantum cell lineages complemented with various HA-tagged PABP1 mutants. (A) PCR results carried out as shown in Fig. 5B evaluating the presence of the endogenous PABP1 gene in recombinant cell lines complemented with different PABP1 mutants. Size markers in bp are shown on the left and the sizes of the amplified bands are also indicated. (B) Western-blot using the serum directed against native PABP1 to assess expression of the non-phosphorylated and phosphorylated PABP isoforms in the cell lines complemented with the wild-type PABP1 gene (HA-PABP1) or the phosphorylation mutant (TP-SP Mut). Notice the extra phosphorylated band in the presence of the wild-type HA-tagged PABP1, which is slightly bigger in comparison with the native PABP1 (L. infan-tum lane) due to the presence of the HA-tag (phosphorylated and non-phosphorylated isoforms indicated by arrows). The experiments shown in (A) and (B) are represen-tative of results generated with a minimum of two sets of independently transfected cells.

PABP1 is not strictly required for cell viability, since the non-phosphorylatable mutant can complement the loss of both copies of the otherwise essential PABP1 gene. Likewise, the protein having the TGM/AAA mutation can also function sufficiently to maintain cell viability in the absence of the endogenous protein. The mutant impaired in its ability to bind to mRNAs and associate with polysomes (YGF/AAA), however, cannot complement the lack of the endogenous PABP1. The same applies for the LMW/AAA mutant, indicating that the LMW motif is involved in critical interactions required for the PABP1 function.

Effects on L. infantum growth induced by the overexpression of PABP1 mutants

The essential nature of the PABP1 function in Leishmania might be linked to a role during translation or to other pro-cesses associated with the metabolism of mRNAs. Since a PABP1 double knockout cell line cannot be generated, an indi-rect way to study its function is to evaluate the effects induced by the overexpression of the native protein or selected mutants in transgenic cells through the transfection with the plasmid encoded genes. Here, to investigate more specifically the effect of the various PABP1 mutations on promastigote and amasti-gote growth, haploid PABP1/PABP1::HYG (HYG SKO) cell lines were used. The SKO cell lines were chosen since the 50% reduction in copy number for the endogenous PABP1 gene might increase the cells' sensitivity to any effects induced by the over-expressed recombinant proteins. Indeed, during the selection phase of the transfection procedure with the plasmids encoding the PABP1 wild-type and mutant proteins, the cells transfected with several of the mutants did take a longer time to recover, when compared with wild-type PABP1 (data not shown). Once recovered, growth of the different cell lines was then compared using standard promastigote growth curves. As described in Fig. 1, these curves were also started with stationary phase promastigotes passaged to new media and allowed to grow for several days, with growth monitored daily. For these experiments cell lines generated from independent transfection events of at least two distinct SKO cell lines were used and the results described were reproduced with a minimum of two independent cell. No major differences in growth were observed, however, when the cells expressing either the wild-type PABP1 or any of the mutants were compared (Fig. 7A). The same cell lines from stationary cultures were then transferred to amastigote media, and allowed to differentiate. Similarly, no significant differences were observed in differentiation and growth as amastigotes for the phosphorylation mutant (TP-SP Mut) as well as the LMW/AAA and TGM/AAA mutants in comparison with the wild-type protein (Fig. 7B). In contrast, cells overexpressing the YGF/ AAA mutant did not grow as amastigotes and this phenotype suggests a dominant negative effect, since the overexpression of the mutant PABP1 might be interfering with the activity of the endogenous protein and inhibiting growth as amastigotes. This effect was not seen with the double mutant protein (LMW/ AAA + YGF/AAA), confirming that the growth inhibitory phe-notype requires an intact LMW motif.

Identification of known and novel PABP1 partners

PABP1 likely executes its functions by binding to poly-A as well as to proteins which are involved in different processes associated to the metabolism of mRNAs, including translation. Leishmania PABP1 has been shown to co-immunoprecipitate with both partners of the EIF4E4/EIF4G3 complex and direct interactions with these two proteins have also been demonstrated [35,49], providing a possible explanation as to how PABP1 would function in translation initiation. Nevertheless, it is not clear what other protein factors or even homologs to eIF4F sub-units might also be interacting with PABP1 and which could be affected by the identified phosphorylation events. Here, in order to investigate the association of PABP1 with known or

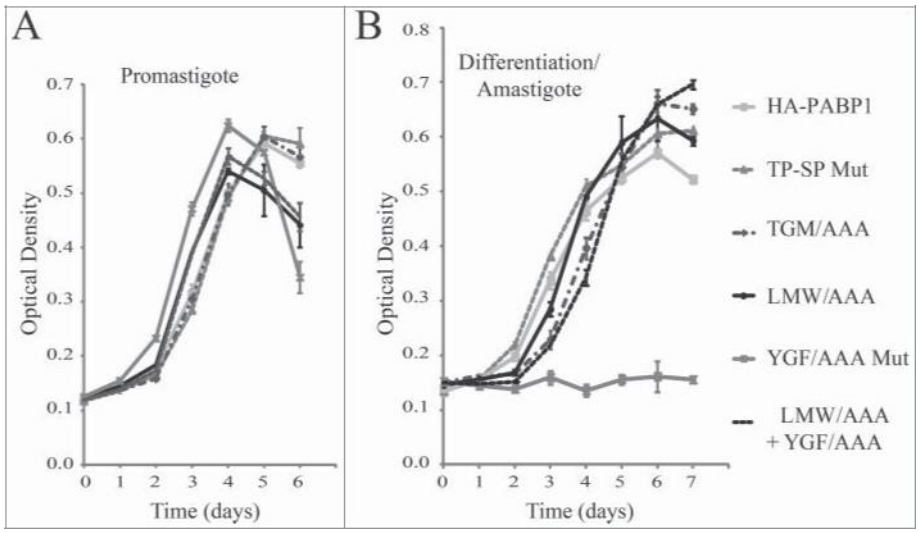


Figure 7. Cell growth phenotype of the haploid PABP1 knockout cell lineages complemented with various HA-tagged PABP1 mutants. (A) and (B) L. infantum cells lacking one of the PABP1 gene copies (SKO) were transfected with episomal vectors expressing either the wild-type HA-PABP1 or the various mutant proteins. Cells were grown as promastigotes (A) or as axenic amastigotes (B) and their growth monitored daily. Standard deviations were derived from a single experiment done with three replicates grown in parallel. Two sets of experiments were done for two sets of independently transfected cells.

novel protein binding partners, cytoplasmic extracts from the cells expressing the HA-tagged wild-type PABP1 were used in immunoprecipitation (IP) assays with an anti-HA monoclonal antibody bound to magnetic beads. Control assays were also carried out with cytoplasmic extracts from non-transfected L. infantum promastigotes incubated with the anti-HA bound magnetic beads and the precipitated samples from both reac-tions were then submitted to mass-spectrometry analysis. Two sets of independent experiments were carried out and, follow-ing polypeptide identification, only those which were substan-tially enriched in the PABP1 sample in comparison with the control, in both experiments, were considered and are listed on Table 1. As expected, the top hit was the HA-tagged PABP1 and both EIF4G3 and EIF4E4 were also found among the poly-peptides bound to PABP1, highlighting the strong association between the Leishmania PABP1 and the EIF4E4/EIF4G3 complex. In contrast, none of the other eIF4E or eIF4G homo-logues were found. The remaining bound polypeptides include the RRM-containing protein RBP23 (LinJ.17.0610), an unchar-acterized protein having both a Nuclear Transport Factor 2 (NTF2) and RRM domains (LinJ.18.0300), a Zinc-finger protein having both KH and Helicase domains (ZC3H41 – LinJ.27.1220) and yet another uncharacterized protein (LinJ.05.0450). The two T. brucei PABP homologues as well as both EIF4E4 and EIF4G3 have been shown to stimulate transla-tion of a reporter mRNA when tethered to its 3 ^{0}UTR [39,40] and the same applies for the T. brucei orthologues of RBP23 and LinJ.18.0300.

The association of several known translational activators with PABP1 suggests a common function involving ribosome recruitment during translation initiation. Parallel IP assays were also carried out using the HA-tagged PABP1 phosphory-lation TP-SP mutant, however no clear differences in co-immu-noprecipitated proteins were observed, when compared with the wild-type protein, which were reproducible in the two experiments performed (data not shown). Since the mutant protein co-exists with the wild-type native PABP1 in the trans-genic cells, and multiple PABP1 molecules may bind to the same mRNA poly-A tail, it is possible that the effects induced by the lack of phosphorylation of the mutant protein may be masked by its association with the native PABP1. Overall, the mass-spectrometry data confirm a strict association between PABP1 and EIF4E4/EIF4G3, but no other eIF4F-like complex or subunits, and identifies novel proteins, which might be required for PABP1 to function properly during translation or other processes.

Table 1. Proteins co-purifying with HA-tagged PABP1 .

Identity	GeneDB ID	Log_2 (I[a] Ratio[b]) 1st Exp	2nd Exp
PABP1	LinJ.35.5360	6.38	5.62
RNA-binding protein – RBP23	LinJ.17.0610	5.59	4.76
EIF4G3	LinJ.16.1700	4.78	3.09
Uncharacterized protein having NTF2 and RRM-like domains	LinJ.18.0300	4.65	4.95
EIF4E4	LinJ.30.0460	4.11	4.18
Zinc-Finger and KH domain containing protein – ZC3H41	LinJ.27.1220	3.67	4.16
Uncharacterized protein	LinJ.05.0450	3.18	3.66

Six polypeptides are shown which specifically co-immunoprecipitated with the HA-tagged PABP1 in two independent immunoprecipitation (IP) experiments, with a minimum of 8-fold increase over the negative control (Log_2 (I Ratio) > 3).

[a]I = Intensity.

[b]Ratio = This was calculated as the ratio between the intensity observed for the IP with extracts derived from cells expressing HA-PABP1 divided by the intensity observed for the extracts from control cells.

Defining the EIF4E4 binding region within PABP1

The motifs within EIF4E4 that are involved in the interaction with PABP1 have been recently identified and seen to be localized within the unique N-terminal segments found only in the trypanosomatid EIF4E3 and EIF4E4 orthologues [52]. However, the binding motifs within PABP1 have not been identified yet and here an attempt was made to better define this interaction. First, truncated PABP1 proteins were expressed as GST-fusions in E. coli and tested for their ability to bind to ^{35}S-labeled EIF4E4 through co-precipitation/pull-down assays. EIF4E4 was seen to bind to the C-terminal half of PABP1 in a segment including the 4th RRM, its linker region and the MLLE domain (C-terminus in Fig. 8A). The PABP1 mutants having the LMW/AAA + YGF/AAA or the TGM/AAA mutations were expressed also as GST-fusions and tested for their ability to interact with labelled EIF4E4. Both wild-type and the LMW/AAA + YGF/AAA full-length proteins bound efficiently to EIF4E4, but no binding was seen with the TGM/AAA mutation, directed to the protein's MLLE domain. These results define then the MLLE domain of PABP1 as the binding site for EIF4E4 (Fig. 8B) with a role for the TGM residues in this interaction.

Molecular modeling of the PABP1/EIF4E4 interaction

It is well known that the MLLE domain found in PABP homologues binds to PAM2 motifs which have been identified in proteins with distinct functions in different organisms, including plants and mammals [33]. The interaction between the MLLE domain and a PAM2 motif has been more recently characterized through X-ray crystallography and the latter has been defined as a region of 12 residues, with five of those (at positions 3, 5, 7, 10 and 12) being directly implicated in the interaction. Binding to PABP requires a critical stacking interaction between the conserved Phe/Tyr within PAM2 at position 10 and another conserved Phe/Tyr localized to the MLLE's second alpha helix [57]. The TGM motif lies within a conserved stretch localized in the middle of the MLLE (see Fig. S1), within its central third helix, and both the threonine and methionine residues within this motif have been implicated in the interaction with PAM2. With the identification of the MLLE domain from PABP1 as the site for its interaction to EIF4E4, and considering that three motifs (boxes A, B and C) have been found in EIF4E4 to be implicated in the binding to PABP1 [52], this interaction was further analyzed. Indeed, when the sequences from the three EIF4E4 boxes were compared with PAM2 motifs found in PABP interacting proteins from various eukaryotes, all three aligned with the consensus motif derived from these sequences (Fig. 9A). Based on this alignment all three boxes can be confidently defined as PAM2 motifs, with box B more closely resembling the consensus motif while boxes A and C are characterized by some divergent residues in positions which are generally conserved in most other PAM2 sequences.

Next, we attempted to model the interaction between the MLLE domain from Leishmania PABP1 with the PAM2 motif found in the EIF4E4 box B. The interaction between the

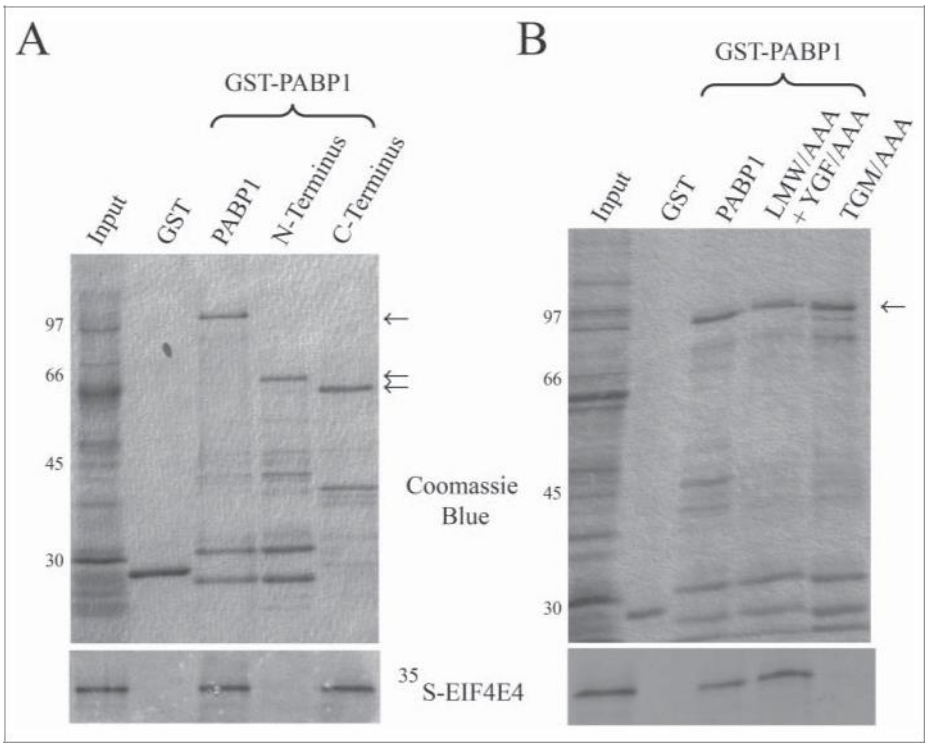


Figure 8. Mapping the EIF4E4 binding region within PABP1. Co-immunoprecipitation assay to map the localization of the EIF4E4 binding motif within PABP1. (A) Recom-binant L. major PABP1 expressed in Escherichia coli fusioned at its N-terminus with Glutathione S-transferase (GST), was assessed for its ability to bind to the [^{35}S]-labeled wild-type EIF4E4. Full-length PABP1 or truncated mutants lacking part of its N- or C-terminal regions (as shown in Fig. 2A) were evaluated for their ability to bind to the labeled EIF4E4. (B) Full-length L. infantum GST-PABP1 was also compared with the "TGM" or the double "LMW"/"YGF" mutants as to their ability to bind to EIF4E4. The upper panels shows the Coomassie-blue stained gels indicating the recombinant GST (negative control), the wild-type GST-PABP1 fusion or the truncated proteins and mutants assayed (the recombinant proteins are indicated by arrows and the sizes of molecular weight markers are shown on the left). The panels below shows the result from the assays carried out evaluating their binding to the wild-type EIF4E4.

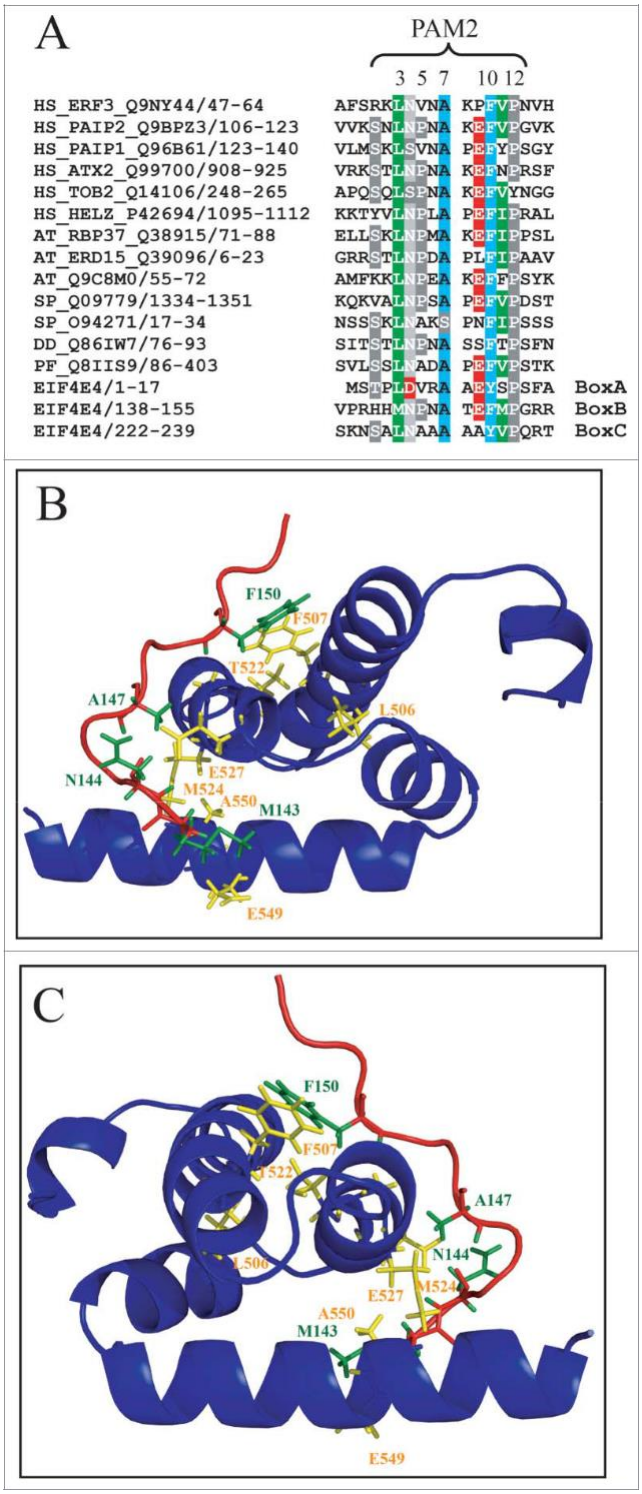


Figure 9. Evaluating the EIF4E4 PAM2 motifs and the EIF4E4/PABP1 interaction. (A) Alignment of multiple PAM2 motifs identified in various putative PABP interacting proteins isolated from different organisms with the three box (A, B and C) motifs reported for L. infantum EIF4E4 [52]. Proteins from human (Homo sapiens – HS), plant (Arabidopsis thaliana – AT), fungi (Schizosaccharomyces pombe – SP) and other unicellular organisms (Dyctiostelium discoideum – DD; Plasmodium falcipa-rum – PF) are shown, as previously described [33]. The positions of the five hydro-phobic residues which characterize the PAM2 motif are shown (3, 5, 7, 10 and 12).
(B) and (C) Modeling of the interaction between the L. infantum EIF4E4 PAM2 motif from box B and the PABP1 MLLE domain. A segment starting at proline 484 and ending in glutamine 588 from the full-length PABP1 was used for the modeling. The alpha-helical segments of the MLLE are in blue with relevant aminoacid resi-dues colored in yellow while the peptide backbone from the PAM2 motif is in red with selected residues colored in green. The image in (C) is horizontally rotated 180 degrees in comparison with (B).

human PABP MLLE and PAM2 motifs from binding partners has been solved by NMR and X-ray crystallography and the MLLE has been characterized as a peptide-binding domain capable of specifically recognizing the PAM2 motif [57,58]. The molecular modeling of the interaction between the Leishmania PABP1 MLLE and the EIF4E4 PAM2 motif is shown in Fig. 9B and 9C, with both the MLLE domain and the PAM2 motif conforming to the published structure. The critical PAM2 residues at positions 3, 4, 7 and 10 (M143, N144, A147 and F150, respectively) are indicated as well as several residues from the MLLE domain involved in the binding between the two polypeptides (also highlighted in Fig. S1). Noteworthy is the stacking interaction between the phenylalanine at position 10 in PAM2 with another phenylalanine residue in MLLE (F507), the single most important event required for the binding to occur [59]. Overall, the modeling analysis strongly supports that the individual residues found in the EIF4E4 box B perform all the expected interactions with the conserved residues found within the PABP1 MLLE domain, compatible with a role as a PAM2 motif, and that the motif on its own would be sufficient to mediate the EIF4E4 interaction with PABP1.

Discussion

The data presented in this work are consistent with phosphorylation of PABP1 being part of a novel pathway in Leishmania species controlling the protein's function and possibly translation. PABP1 is phosphorylated simultaneously with its partner EIF4E4 (this work and [52]) and the similarities in the targeted residues are compatible with both proteins being targeted by the same kinase(s). Indeed, both T. brucei orthologues of PABP1 and EIF4E4, as well as a few unrelated proteins, were found to be targeted by the cell cycle dependent kinase CRK1, in complex with the cyclin CYC2 [60]. CRK1 is conserved in Leishmania and is involved in controlling the G1/S transition [61], so phosphorylation of both PABP1 and EIF4E4 might be required for the cells to proceed through the cell cycle. Never-theless, the phosphorylated pattern seen by both proteins soon after passaging of stationary Leishmania promastigotes to new media also indicates a phosphorylation mediated by MAP-kinase like enzymes, which yet need to be identified. Phosphor-ylation of PABP1 is not needed for its association to polysomes and is not required for the protein to function and maintain cell viability. In vitro the strong association between PABP1 and EIF4E4 does not require phosphorylation and the efficient phosphorylation of the TGM/AAA mutant implies that binding of the two proteins might not be required for phosphorylation to occur. Likewise, the efficient phosphorylation of the YGF/
AAA mutant, impaired in its association to polysomes and presumably to mRNAs, indicate that PABP1 does not necessarily need to be tightly bound to mRNAs to be phosphorylated, although the lack of phosphorylation by the double mutant (LMW/AAA + YGF/AAA) does indicate that the protein needs to retain some functionality in order to be properly phosphory-lated. The fact that both binding partners, EIF4E4 and PABP1, are simultaneously phosphorylated is a strong indication of an impact on translation that needs to be further resolved.

In different organisms, PABP phosphorylation has been seen to affect many of its interactions and functions. In plants, for example, PABP hypophosphorylation lowers its binding activity to the poly(A) and eIF4B while its interaction with eIF4G is enhanced [62]. In Xenopus oocytes, hyperphosphorylation of an embryo homologue of PABP, ePABP, also targets the protein's linker region but the phosphorylated serine or threonine residues do not require a neighboring proline, as seen for the Leishmania PABP1. The ePABP phosphorylation is dispensable for translational activation, but it is required for the cytoplasmic polyadenylation of maternal mRNAs [63]. In HeLa cells, PABP phosphorylation has also been confirmed [64], although subsequently PABP1 was shown to be targeted by multiple post-translational modifications events, which include aspartate and glutamate methylation, lysine and arginine methylation and lysine acetylation in a report where no PABP phosphorylation was seen [65]. It is likely then that different organisms evolved alternative mechanisms of post-translationally modifying PABP homologues in order to regulate its function.

The solution structure of the C-terminal domain from the T. cruzi PABP2 confirms the conservation of the MLLE domain from trypanosomatid PABP homologues and its ability to bind PAM2-like motifs [66]. PAM2 motifs have been shown to localize to disordered regions mapped near to protein phosphorylation sites [67] and some PABP partners have been found to have two PAMs, one of which is responsible for specific PABP recognition, while the other might be required to stabilize the interaction [68,69]. All these features seem to be conserved within the motifs found in the Leishmania EIF4E4 N-terminus and support the in vitro binding assays and bioinformatic analysis, which have defined the PABP1/EIF4E4 inter-action. The characterization of this interaction highlights even further the diversity observed in PABP interacting proteins which bind through the PAM2/MLLE interaction. This has been more recently discussed in plants, where the large number of PABP homologues is probably associated with multiple pro-cesses involved in different aspects of mRNA metabolism and which might have major impacts on regulation of gene expres-sion and translation [34,70].

It is still not clear what role phosphorylation might have in regulating both PABP1 and EIF4E4 functions and presumably leading to enhanced translation of the bound mRNAs under conditions of logarithmic cell growth. Human proteins have been identified having PAM2 motifs similar to those observed for EIF4E4 and localized next to phosphorylation sites within disordered regions. Phosphorylation of these proteins inhibits the interaction with PABP [16,67], therefore it seems likely that simultaneous phosphorylation of both PABP1 and EIF4E4 would inhibit or remodel their interaction, something that needs to be confirmed. Evaluating this possibility, however, has so far been prevented by the strong susceptibility of both proteins to dephosphorylation, and degradation in the case of EIF4E4, during cytoplasmic extract preparation. The lack of a clear phenotype by either the PABP1 or EIF4E4 phosphorylation mutants in the complementation assays (this work and [52]) also prevented a proper evaluation of the role of phosphorylation in translation. This might be due to the fact that when one protein is mutated the other is still phosphorylated and both might be required to be in a dephosphorylated state in order to induce differences in growth profile or protein synthesis, so both proteins would have to be targeted simultaneously. Further work then will be required in order to investigate these issues properly.

The identification of protein partners which specifically co-immunoprecipitated with PABP1 also helped clarify its func-tional role. PABP1 was found to interact with EIF4E4/EIF4G3, constituting the eIF4F complex most likely involved in transla-tion initiation in Leishmania [44], and none of the other eIF4E or eIF4G homologues identified in this organism. Likewise, nei-ther PABP2 or PABP3 seem to co-precipitate with the EIF4E4/ EIF4G3 complex [35], highlighting the strong association between PABP1 and EIF4E4/EIF4G3 with likely common roles in translation initiation. Both RBP23 and the uncharacterized LinJ.18.0300 polypeptide are likely RNA binding proteins, which may help recruiting PABP1 and the EIF4E4/EIF4G3 complex to selected mRNAs in order to promote their translation. The other two putative PABP1 partners, the CCCH zinc finger protein ZC3H41 and the uncharacterized LinJ.05.0450, might be involved in regulating PABP1 function either in translation or otherwise. In mammals PABP interacting proteins (Paip1 and Paip2) are well known translation regulators which function by directly binding to PABP, either enhancing or repressing translation [71]. More recently, in T. brucei, another CCCH zinc finger protein, ZC3H11, has been shown to act as a post-transcriptional regula-tor required to stabilize stress response mRNAs following heat shock [72]. It is possible that ZC3H41 also regulates translation acting upon selected messages or otherwise functions in other processes associated with the metabolism of mRNAs. So far how-ever, and in view of the fact that PABP2, but not PABP1, can shuttle to the nucleus upon inhibition of mRNA synthesis or processing [35,36], it is more likely that PABP1 and binding part-ners would not be involved in processes such as mRNA process-ing and transport, which would require a nuclear localization.

Materials and methods

Parasite growth and transfections

Leishmania infantum MHOM/MA/67/ITMAP-263 promastigotes were cultured in SDM-79 medium supplemented with 10% heat-inactivated FCS (Multicell Wisent Inc.) and 5 mg/ml hemin and axenic amastigotes in MAA/20 medium supplemented with 20% FCS in 5% CO_2 atmosphere as described previously [52]. Growth curves were set up using late-stationary phase cells passaged to fresh medium at a pre-established cell density of 10^6 cells/ml. Samples were taken at selected time points for processing, SDS-PAGE and blotting. Transfection procedures used for both circu-lar plasmids (episomal expression) and linear DNA fragments (for integration and gene deletion) were carried out by electropo-ration as described [73]. Cells transfected with the PABP1 wild-type and mutant constructs were selected with neomycin (G418, 20 mg/ml, Sigma) and those transfected with the linear DNA cas-settes for PABP1 gene deletion events with hygromycin B (40 mg/ ml, Sigma) or puromycin (70 mg/ml, Sigma).

Antibodies and western-blotting

Western-blotting was performed using standard procedures. Native Leishmania PABP1 and PABP2 were detected using

rabbit polyclonal sera raised against the two proteins as described previously [35]. The anti-mouse alpha-tubulin antibody (1:10000 dilution; Sigma) was used for normalizing protein loading. For the detection of the different HA-tagged proteins, an anti-HA monoclonal antibody (100 hg/ml) from Applied Biological Materials was used. For the Leish-mania large ribosomal P0 subunit (TriTrypDb accession: LmjF.27.1380), antiserum was raised by immunizing adult New Zealand White rabbits with a His-tagged recombinant protein. Recombinant P0 was produced after amplification of the L. major gene flanked by restriction sites for BamHI and XhoI immediately before and after the AUG and stop codons, respec-tively (oligonucleotides listed in Table S1). The amplified frag-ment was cloned into the same sites of the pET21a vector (Novagen), followed by transformation into Escherichia coli BL21 cells, protein expression and affinity purification using Ni-NTA Agarose (Qiagen).

Sucrose gradient methods

Cytoplasmic cellular extracts were produced from late-logarithmically grown L. infantum promastigotes incubated for 10 minutes with 100 mg/ml cycloheximide (Sigma). For extract preparation, the cells were first harvested and washed once in PBS and once again in Hepes-lysis buffer (20 mM Hepes-KOH, pH7.4, 75 mM potassium acetate, 4 mM magnesium acetate, 2 mM DTT), both supplemented with 1x protease inhibitors (EDTA-free EASYpack Roche) and 100 mg/ml cycloheximide. The weight of the final pelleted cells was measured prior to their resuspension to a concentration of 2–4 £ 10^9 cells/ml in Hepes-lysis buffer plus protease inhibitors and cycloheximide. Twice the weight of the pelleted cells in glass beads, 150–200 mm in diameter (from Sigma) was then added. Lysis was car-ried through vortexing for 15 minutes at 4 C, followed by pas-saging five times through a 30 gauge needle and centrifugation at 13,000g for 15 minutes to remove cellular debris. Roughly 15 to 20 $OD_{260\ nm}$ units of the supernatants, the cytoplasmic extracts, were then layered on top of 15% to 45% linear sucrose gradients. For the EDTA-treated control gradients, 0.1 M EDTA was added to the cytoplasmic extracts prior to loading on the gradients. Sucrose gradient preparations and fraction-ation were carried out as described previously [74,75] and ali-quots from each fraction were prepared for SDS-PAGE and western-blotting through 10% TCA precipitation followed by acetone washes and resuspension in SDS-PAGE sample buffer.

Sequence analysis and molecular modeling

Sequence analysis and alignment of the various PABP1 homo-logues were carried out as previously described [52] using selected trypanosomatid sequences recovered from TriTrypDB and sequences from other organisms from GenBank. TriTrypDB accessions: Leishmania infantum (Li) PABP1 – LinJ.35.5360; LiPABP2 – LinJ.35.4200; LiPABP3 – LinJ.25.0080; L. braziliensis (Lb) PABP1 – LbrM.34.4980; LbPABP2 – LbrM.34.4130; LbPABP3 – LbrM.25.0080; Trypanosoma brucei (Tb) PABP1 – Tb927.9.9290; TbPABP2 – Tb927.9.10770. Further GeneBank accessions: Homo sapiens (Hs) PABP1 – NP_002559.2; Saccharo-myces cerevisiae (Sc) PABP – NP_011092; Arabidopsis thaliana (At) PABP2 – NP_195137 (the most conserved of the various A. thaliana PABPs). The alignment of the various PAM2 motifs was carried out using similar procedures but using colored shad-ings when more than 60% of the sequences were identical or sim-ilar. All PAM2 sequences used in the alignment have been previously described [33,52].

Modeling of the MLLE domain from L. infantum PABP1 was performed using the Modeller program [76], version 9.14, with the crystal structure from the human PABP MLLE domain bound to a PAM2 motif from the PABP partner PAIP2 [57] used as template and downloaded from PDB (http://www.rcsb.org/pdb/home/home.do, PDB ID: 3KUS). The PyMOL soft-ware, version 1.7 (PyMOL Molecular Graphics System, Version 1.7 Schr€odinger, LLC), was then applied to overlap the best model produced with the original crystallized structure of the human PABP MLLE bound to the PAIP2 PAM2 motif. The next step was the manual removal of the crystallized MLLE domain, resulting in a model consisting of the L. infantum PABP1 MLLE domain bound to the PAM2 motif from human PAIP2. To replace the original PAM2 motif with the one mapped to the N-terminus from L. infantum EIF4E4, tools from the Rosetta package [77] were used. First, fixbb was used to replace residues non-conserved between the crystallized PAM2 motif and the selected motif from EIF4E4. This replace-ment does not allow movement for the backbone of the PAM2 motif, which may produce unstable final conformations. The flexpepdock tool, also from the Rosetta package, was thus applied with 300 solutions in order to make the PAM2 back-bone flexible and better accommodate the EIF4E4 motif on the MLLE domain model from the L. infantum PABP1. The final models were assessed by PROCHECK [78], version 3.5, which validated them to be used to infer the interaction between the PABP1 MLLE domain and the EIF4E4 PAM2 motif.

Plasmid constructs, DNA manipulations

L. infantum genomic DNA extraction and PCR amplification were carried out as described previously [52]. The full length wild-type PABP1 gene was amplified together with flanked BamHI and HindIII restriction sites and a 27 nucleotide exten-sion encoding a single copy of the HA epitope (YPYDVPDYA) added immediately prior to the PABP1 translation stop codon. Site-directed mutagenesis was also carried out as described [52]. All PCR fragments were first cloned into the pGEM-T Easy vector (Promega) prior to sequencing and subcloning into the BamHI-HindIII sites of the Leishmania expression vector pSPBT1YNEOa [79] with the 3 ^{0} HA tag. The PABP1 gene dele-tion and complementation strategies followed the same proce-dures used previously with the L. infantum EIF4E4 gene [52]. All oligonucleotides used for the various amplification and mutagenesis reactions are listed in Tables S2 through S4. The genes encoding selected L. infantum PABP1 mutants were sub-sequently subcloned into the BamHI-HindIII sites of a modi-fied pGEX4T3 vector (GE Healthcare) for the expression of GST-tagged recombinant proteins. The two truncated versions of the L. major GST-tagged PABP1 (only five amino acid differ-ences between the L. infantum and L. major orthologues) were generated from the wild-type PABP1 cloned into the pGEX4T3 vector [35]. For the N-terminus, the plasmid was digested with

BamHI/NotI and the 897 pb fragment encoding for the first three PABP1 RRMs (residues 1 to 313 from the wild type protein) was gel-purified and subcloned into the same sites of pGEX4T3. For the C-terminus, a gene fragment encoding residues 302 to 560 was amplified flanked by sites for NcoI/HindIII, purified and cloned likewise into the same sites of a modified pGEX4T3 vector. Southern blots were carried out using standard procedures after total DNA digestion with ClaI or PvuII. The probe used was derived from the 500 bp PCR fragment corresponding to the 5 ^{0} intergenic region immediately before the PABP1 gene start codon and generated using the primers listed in Tables S3.

Extract preparation for immunoprecipitation

Whole cytoplasmic extracts from wild-type L. infantum or recombinant strains expressing selected HA-tagged PABP1 variants were generated after lysing the cells through cavitation, aiming to reduce undesired debris and rupture of the nuclei and/or organelles. First, late-logarithmically grown Leishmania infantum promastigotes were harvested and washed once in ice cold PBS, followed by resuspension in HEPES-lysis buffer (20 mM HEPES-KOH, pH 7.4, 75 mM potassium acetate, 4 mM magnesium acetate, 2 mM DTT, supplemented with and 1x EDTA-free protease inhibitors from Roche) to a concentration of 1–2 £ 10^{9} cells/ml. Lysis was carried out using nitrogen cavitation as described [80], with modifications. Briefly, the resuspended cells were trans-ferred into the cavitation chamber of the cell disruption vessel (Parr Instruments) and incubated at 4 C under 70 bar pres-sure for 40 min, followed by rapid decompression and lysis. The lysates were submitted to centrifugation at 17,000g for 10 min to remove cellular debris and the supernatants, the cytoplasmic extracts, aliquoted and stored.

For the immunoprecipitation assays (IPs), the cytoplasmic extracts were mixed with Pierce^{TM} Anti-HA Magnetic Beads as per manufacturer's protocol. Briefly, 0.2 mg of these anti-HA magnetic beads were washed three times with PBS followed by the incubation with the cytoplasmic extract for 1 h at 4 C. After, the depleted supernatant was removed and the beads were washed three times with PBS, the resulting, specifically bound, immunoprecipitated antigen-antibody complexes were eluted in SDS-PAGE sample buffer. They were then assayed by SDS-PAGE and Western-blotting using antibodies against PABP1 and the HA-tag to confirm the efficiency of the precipitation reaction.

Mass spectrometry analysis

For mass spectrometry (MS), the sets of eluted proteins were loaded onto 15% SDS-PAGE gels and allowed to migrate into the resolving gel. Gel slices containing the whole IP products were then excised and submitted to an in-gel tryptic digestion and mass spectrometry analysis as previously described [81]. Protein identification was based on the L. infantum protein sequence database downloaded from UniProt. For validation of the identified proteins a minimum of six amino acids for peptide length and two peptides per protein were required. In addition, a false discovery rate (FDR) threshold of 0.01 (using the decoy database approach) was applied at both peptide and protein levels. To confirm the specificity of the IP assays for each polypeptide, the ratio between the intensity generated from the IPs using the extracts expressing the HA-tagged PABP1 and the intensity from the control IP using an extract from non-transfected cells was first determined. The base 2 logarithms of the values produced were then calcu-lated for two independent experiments carried out with dif-ferent cytoplasmic extracts and only the values >3.0 were considered. Identified polypeptides were then ranked and listed in Table 1 with the highest ranking values obtained for the first experiment listed on top.

In vitro pull-down assays

Co-precipitation/pull-down assays were essentially performed as described previously [42,52] using Glutathione-Sepharose 4B beads (GE Healthcare) and affinity purified GST-tagged recombinant proteins. GST alone and the different GST-tagged PABP1 variants (wild-type, truncations and mutants) were expressed in Escherichia coli, immobilized on the beads and assayed for their ability to bind to ^{35}S-labeled EIF4E4. The labeled protein was obtained through the linearization of the corresponding pET21a derived plasmids with NotI, followed by transcription with T7 RNA polymerase in the presence of the cap analogue and translation in rabbit reticulocyte lysates (Promega or Ambion) supplemented with ^{35}S-methionine (Perkin Elmer).

Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

Funding

CNPq - Brazil, 480899/2013-4 Capes - Brazil, 23038.007656/2011-92 CNPq - Brazil, 313934/2013-4 Facepe - Brazil, APQ-0239-2.02/12 MDEIE - Canada, PSR-SIIRI-439 NSERC - Canada, 418444 Government of Can-ada j Canadian Institutes of Health Research (CIHR), MOP-12182.

Acknowledgments

We thank Carole Dumas for technical assistance and other members of Dr. Papadopoulou's and Dr. de Melo Neto's laboratories for helpful discussions. We are also thankful to the Proteomics platform of the CHU de Quebec Research Center for preliminary mass-spectrometry analysis of PABP1 bound complexes. The work in Dr. Papadopoulou's lab was supported by the Canadian Institutes of Health Research (CIHR) operating grant MOP-12182, the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant GC100741 (418444), and the Ministere du Developpement

Economique de l Innovation et de l Exportation (MDEIE) PSR-SIIRI-439 grant awarded to BP. The Leishmania work in Dr. de Melo Neto's lab was more recently funded with grants provided by the Brazilian funding agencies FACEPE (APQ-0239-2.02/12 and APQ-1662-2.02/15), CNPq (480899/2013-4, 313934/2013-4 and 401282/2014-7) and CAPES (23038.007656/2011-92). Dr. de Melo Neto was the recipient of a Visiting Scholar grant from CAPES for a medium term stay at the University of Laval (BEX 3743/10-1). To cover their stays at the same institution, TPR, TDCL and LMN received funding from CNPq, the Agence Universitaire de la Francophonie and CAPES, respectively. Studentships for the graduate students in Brazil (TPR, KCM, POR, LAA, LMN and CCX) were provided by CNPq, FACEPE or CAPES. The authors thank the Nucleo de Plataformas Tecnologicas (NPT) at the Instituto Aggeu Magalh~aes for the use of its automatic sequencing facility,

the Instituto Carlos Chagas for the use of its proteomics facility and FIOC-RUZ for the funding support.

References

[1] Ouellette M, Papadopoulou B. Coordinated gene expression by post-transcriptional regulons in African trypanosomes. J Biol. 2009;8:100. doi:10.1186/jbiol203

[2] Fernandez-Moya SM, Estevez AM. Posttranscriptional control and the role of RNA-binding proteins in gene regulation in trypanosoma-tid protozoan parasites. Wiley Interdiscip Rev RNA. 2010;1:34–46. doi:10.1002/wrna.6

[3] De Gaudenzi JG, Noe G, Campo VA, Frasch AC, et al. Gene expres-sion regulation in trypanosomatids. Essays Biochem. 2011;51:31–46. doi:10.1042/bse0510031

[4] Schwede A, Kramer S, Carrington M. How do trypanosomes change gene expression in response to the environment? Protoplasma. 2012;249:223–38. doi:10.1007/s00709-011-0282-5

[5] Kramer S. Developmental regulation of gene expression in the absence of transcriptional control: the case of kinetoplastids. Mol Biochem Parasitol. 2012;181:61–72. doi:10.1016/j.molbiopara.2011.10.002

[6] Clayton CE. Networks of gene expression regulation in Trypanosoma brucei. Mol Biochem Parasitol. 2014;195:96–106. doi:10.1016/j. molbiopara.2014.06.005

[7] Clayton CE. Gene expression in Kinetoplastids. Curr Opin Micro-biol. 2016;32:46–51. doi:10.1016/j.mib.2016.04.018

[8] van der Kelen K, Beyaert R, Inze D, et al. Translational control of eukaryotic gene expression Crit Rev Biochem Mol Biol. 2009;44:143–68. doi:10.1080/10409230902882090

[9] Gingras AC, Raught B, Sonenberg N. eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. Annu Rev Biochem. 1999;68:913–63. doi:10.1146/annurev.biochem.68.1.913

[10] Gallie DR. The role of the poly (A) binding protein in the assembly of the cap-binding complex during translation initiation in plants. Translation. 2014;731:1–14

[11] Merrick WC. eIF4F: A Retrospective. J Biol Chem. 2015;290:24091–9. doi:10.1074/jbc.R115.675280

[12] Aylett CHS, Ban N. Eukaryotic aspects of translation initiation brought into focus. Philos Trans R Soc London B Biol Sci. 2017;372: doi:10.1098/rstb.2016.0186

[13] Kuhn€ U, Wahle E. Structure and function of poly(A) binding pro-teins. Biochim Biophys Acta. 2004;1678:67–84. doi:10.1016/j. bbaexp.2004.03.008

[14] Gorgoni B, Gray NK. The roles of cytoplasmic poly(A)-binding pro-teins in regulating gene expression: a developmental perspective. Brief Funct Genomic Proteomic. 2004;3:125–41. doi:10.1093/bfgp/3.2.125

[15] Lemay J-F, Lemieux C, St-Andre O, et al. Crossing the borders: Poly (A)-binding proteins working on both sides of the fence. RNA Biol. 2010;7:291–5. doi:10.4161/rna.7.3.11649

[16] Eliseeva IA, Lyabin DN, Ovchinnikov LP. Poly(A)-binding proteins: structure, domain organization, and activity regulation. Biochem. 2013;78:1377–91

[17] Smith RWP, Blee TKP, Gray NK. Poly(A)-binding proteins are required for diverse biological processes in metazoans. Biochem Soc Trans. 2014;42:1229–37. doi:10.1042/BST20140111

[18] Wigington CP, Williams KR, Meers MP, et al. Poly(A) RNA-binding proteins and polyadenosine RNA: new members and novel functions. Wiley Interdiscip Rev RNA. 2014;13:601–22. doi:10.1002/wrna.1233

[19] Gray NK, Hrabalkova L, Scanlon JP, et al. Poly(A)-binding proteins and mRNA localization: who rules the roost? Biochem Soc Trans. 2015;43:1277–84. doi:10.1042/BST20150171

[20] Goss DJ, Kleiman FE. Poly(A) binding proteins: are they all created equal? Wiley Interdiscip Rev RNA. 2013;4:167–79. doi:10.1002/wrna.1151

[21] Burd CG, Dreyfuss G. Conserved structures and diversity of functions of RNA-binding proteins. Science. 1994;265:615–21. doi:10.1126/science.8036511

[22] Deo RC, Bonanno JB, Sonenberg N, et al. Recognition of polyadeny-late RNA by the poly(A)-binding protein. Cell. 1999;98:835–45. doi:10.1016/S0092-8674(00)81517-2

[23] Kuhn U, Pieler T, Kuhn€ U, et al. Xenopus poly(A) binding protein: functional domains in RNA binding and protein-protein interaction. J Mol Biol. 1996;256:20–30. doi:10.1006/jmbi.1996.0065

[24] Kessler SH, Sachs AB. RNA recognition motif 2 of yeast Pab1p is required for its functional interaction with eukaryotic translation initiation factor 4G. Mol Cell Biol. 1998;18:51–7. doi:10.1128/MCB.18.1.51

[25] Imataka H, Gradi A, Sonenberg N. A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. EMBO J. 1998;17:7480–9. doi:10.1093/emboj/17.24.7480

[26] Cheng S, Gallie DR. eIF4G, eIFiso4G, and eIF4B bind the poly(A)-binding protein through overlapping sites within the RNA recognition motif domains. J Biol Chem. 2007;282:25247–58. doi:10.1074/jbc.M702193200

[27] Khanam T, Muddashetty RS, Kahvejian A, et al. Poly(A)-binding protein binds to A-rich sequences via RNA-binding domains 1+2 and 3+4. RNA Biol. 2006;3:170–7. doi:10.4161/rna.3.4.4075

[28] Sladic RT, Lagnado CA, Bagley CJ, et al. Human PABP binds AU-rich RNA via RNA-binding domains 3 and 4. Eur J Biochem. 2004;271:450–7. doi:10.1046/j.1432-1033.2003.03945.x

[29] Khacho M, Mekhail K, Pilon-Larose K, et al. eEF1A Is a Novel Com-ponent of the Mammalian Nuclear Protein Export Machinery. Mol Biol Cell. 2008;19:5296–308. doi:10.1091/mbc.E08-06-0562

[30] Melo EO, Dhalia R, Martins de SC, et al. Identification of a C-termi-nal poly(A)-binding protein (PABP)-PABP interaction domain: role in cooperative binding to poly (A) and efficient cap distal transla-tional repression. J Biol Chem. 2003;278:46357–68. doi:10.1074/jbc. M307624200

[31] Lin J, Fabian M, Sonenberg N, et al. Nanopore detachment kinetics of poly(A) binding proteins from RNA molecules reveals the critical role of C-terminus interactions. Biophys J. 2012;102:1427–34. doi:10.1016/j.bpj.2012.02.025

[32] Khaleghpour K, Kahvejian A, De CG, Roy G, et al. Dual interactions of the translational repressor Paip2 with poly(A) binding protein. Mol Cell Biol. 2001;21:5200–13. doi:10.1128/MCB.21.15.5200-5213.2001

[33] Albrecht M, Lengauer T. Survey on the PABC recognition motif PAM2. Biochem Biophys Res Commun. 2004;316:129–38. doi:10.1016/j.bbrc.2004.02.024

[34] Jimenez-Lopez D, Bravo J, Guzman P. Evolutionary history exposes radical diversification among classes of interaction partners of the MLLE domain of plant poly(A)-binding proteins. BMC Evol Biol. 2015;15:195. doi:10.1186/s12862-015-0475-1

[35] da Costa Lima TD, Moura DMN, Reis CRS, et al. Functional charac-terization of three leishmania poly(a) binding protein homologues with distinct binding properties to RNA and protein partners. Eukar-yot Cell. 2010;9:1484–94. doi:10.1128/EC.00148-10

[36] Kramer S, Bannerman-Chukualim B, Ellis L, et al. Differential locali-zation of the two T. brucei poly(A) binding proteins to the nucleus and RNP granules suggests binding to distinct mRNA pools. PLoS One. 2013;8:e54004. doi:10.1371/journal.pone.0054004

[37] Pitula J, Ruyechan WT, Williams N. Trypanosoma brucei: identifica-tion and purification of a poly(A)-binding protein. Exp Parasitol. 1998;88:157–60. doi:10.1006/expr.1998.4211

[38] Guerra N, Vega-Sendino M, Perez-Morgado MI, et al. Identifica-tion and functional characterization of a poly(A)-binding protein from Leishmania infantum (LiPABP). FEBS Lett. 2011;585:193–8. doi:10.1016/j.febslet.2010.11.042

[39] Erben ED, Fadda A, Lueong S, et al. A genome-wide tethering screen reveals novel potential post-transcriptional regulators in Trypanosoma brucei. PLoS Pathog. 2014;10:e1004178. doi:10.1371/journal.ppat.1004178

[40] Lueong S, Merce C, Fischer B, et al. Gene expression regulatory net-works in Trypanosoma brucei: Insights into the role of the mRNA-binding proteome. Mol Microbiol. 2016;100:457–71. doi:10.1111/ mmi.13328

[41] Klein C, Terrao M, Inchaustegui GD, et al. Polysomes of Trypa-nosoma brucei: Association with Initiation Factors and RNA-Binding Proteins. PLoS One. 2015;10:e0135973. doi:10.1371/journal.pone.0135973

[42] Dhalia R, Reis CR, Freire ER, et al. Translation initiation in Leishmania major: characterisation of multiple eIF4F subunit homologues. Mol Biochem Parasitol. 2005;140:23–41. doi:10.1016/j.molbiopara.2004.12.001

[43] Jagus R, Bachvaroff TR, Joshi B, et al. Diversity of eukaryotic transla-tional initiation factor eIF4E in protists. Comp Funct Genomics. 2012;2012:134839. doi:10.1155/2012/134839

[44] de Melo Neto OP, Reis CR, Moura DM, Freire ER, et al. Unique and conserved features of the protein synthesis apparatus in trypanoso-matid (Trypanosoma and Leishmania) species. In: Hernandez G, Jagus R, editors. Evolution of the protein synthesis machinery and its regulation. Springer International Publishing; 2016. p. page 435–75

[45] Zinoviev A, Shapira M. Evolutionary conservation and diversification of the translation initiation apparatus in trypanosomatids. Comp Funct Genomics. 2012;2012:813718. doi:10.1155/2012/813718

[46] Freire E, Sturm N, Campbell D, et al. The role of cytoplasmic mRNA cap-binding protein complexes in Trypanosoma brucei and other trypanosomatids. Pathogens. 2017;6:55. doi:10.3390/pathogens6040055

[47] Yoffe Y, Leger M, Zinoviev A, et al. Evolutionary changes in the Leishmania eIF4F complex involve variations in the eIF4E-eIF4G interactions. Nucleic Acids Res. 2009;37:3243–53. doi:10.1093/nar/gkp190

[48] Freire ER, Dhalia R, Moura DMN, et al. The four trypanosomatid eIF4E homologues fall into two separate groups, with distinct features in primary sequence and biological properties. Mol Biochem Parasitol. 2011;176:25–36. doi:10.1016/j.molbiopara.2010.11.011

[49] Zinoviev A, Leger M, Wagner G, et al. A novel 4E-interacting protein in Leishmania is involved in stage-specific translation pathways. Nucleic Acids Res. 2011;39:8404–15. doi:10.1093/nar/gkr555

[50] Zinoviev A, Manor S, Shapira M. Nutritional stress affects an atypical cap-binding protein in Leishmania. RNA Biol. 2012;9:1450–60. doi:10.4161/rna.22709

[51] Moura DMN, Reis CRS, Xavier CC, et al. Two related trypanosoma-tid eIF4G homologues have functional differences compatible with distinct roles during translation initiation. RNA Biol. 2015;12:305– 19. doi:10.1080/15476286.2015.1017233

[52] de Melo Neto OP, da Costa Lima TDC, Xavier CC, et al. The unique Leishmania EIF4E4 N-terminus is a target for multiple phosphorylation events and participates in critical interactions required for translation initiation. RNA Biol. 2015;12:1209–21. doi:10.1080/15476286.2015.1086865

[53] Bates EJ, Knuepfer E, Smith DF. Poly(A)-binding protein I of Leishmania: functional analysis and localisation in trypanosomatid parasites. Nucleic Acids Res. 2000;28:1211–20. doi:10.1093/nar/28.5.1211

[54] Urbaniak MD, Martin DM, Ferguson MA. Global quantitative SILAC phosphoproteomics reveals differential phosphorylation is widespread between the procyclic and bloodstream form lifecycle stages of Trypanosoma brucei. J Proteome Res. 2013;12:2233–44. doi:10.1021/pr400086y

[55] Pereira MMC, Malvezzi AM, Nascimento LM, et al. The eIF4E subu-nits of two distinct trypanosomatid eIF4F complexes are subjected to differential post-translational modifications associated to distinct growth phases in culture. Mol Biochem Parasitol. 2013;190:82–6. doi:10.1016/j.molbiopara.2013.06.008

[56] Pyronnet S, Imataka H, Gingras AC, et al. Human eukaryotic translation initiation factor 4G (eIF4G) recruits mnk1 to phosphorylate eIF4E. EMBO J. 1999;18:270–9. doi:10.1093/emboj/18.1.270

[57] Kozlov G, Menade M, Rosenauer A, et al. Molecular determinants of PAM2 recognition by the MLLE domain of poly(A)-binding protein. J Mol Biol. 2010;397:397–407. doi:10.1016/j.jmb.2010.01.032

[58] Kozlov G, De CG, Lim NS, Siddiqui N, et al. Structural basis of ligand recognition by PABC, a highly specific peptide-binding domain found in poly(A)-binding protein and a HECT ubiquitin ligase. EMBO J. 2004;23:272–81. doi:10.1038/sj.emboj.7600048

[59] Kozlov G, Gehring K. Molecular basis of eRF3 recognition by the MLLE domain of poly(A)-binding protein. PLoS One. 2010;5:e10169. doi:10.1371/journal.pone.0010169

[60] Hu H, Gourguechon S, Wang CC, et al. The G1 cyclin-dependent kinase CRK1 in Trypanosoma brucei regulates anterograde protein transport by phosphorylating the COPII subunit Sec31. J Biol Chem. 2016;291:15527–39. doi:10.1074/jbc.M116.715185

[61] Naula C, Parsons M, Mottram JC. Protein kinases as drug tar-gets in trypanosomes and Leishmania. Biochim Biophys Acta. 2005;1754:151–9. doi:10.1016/j.bbapap.2005.08.018

[62] Le H, Browning KS, Gallie DR. The phosphorylation state of poly(A)-binding protein specifies its binding to poly(A) RNA and its interaction with eukaryotic initiation factor (eIF) 4F, eIFiso4F, and eIF4B. J Biol Chem. 2000;275:17452–62. doi:10.1074/jbc.M001186200

[63] Friend K, Brook M, Bezirci FB, et al. Embryonic poly(A)-bind-ing protein (ePAB) phosphorylation is required for Xenopus oocyte maturation. Biochem J. 2012;445:93–100. doi:10.1042/BJ20120304

[64] Ma S, Musa T, Bag J. Reduced stability of mitogen-activated protein kinase kinase-2 mRNA and phosphorylation of Poly(A)-binding Protein (PABP) in cells overexpressing PABP. J Biol Chem. 2006;281:3145–56. doi:10.1074/jbc.M508937200

[65] Brook M, McCracken L, Reddington JP, et al. The multifunctional poly(A)-binding protein (PABP) 1 is subject to extensive dynamic post-translational modification, which molecular modelling suggests plays an important role in co-ordinating its activities. Biochem J. 2012;441:803–12. doi:10.1042/BJ20111474

[66] Siddiqui N, Kozlov G, D'Orso I, et al. Solution structure of the C-ter-minal domain from poly(A)-binding protein in Trypanosoma cruzi: a vegetal PABC domain. Protein Sci. 2003;12:1925–33. doi:10.1110/ ps.0390103

[67] Huang K-L, Chadee AB, Chen C-Y a, et al. Phosphorylation at intrinsically disordered regions of PAM2 motif-containing proteins modulates their interactions with PABPC1 and influ-ences mRNA fate. RNA. 2013;19:295–305. doi:10.1261/ rna.037317.112

[68] Huntzinger E, Braun JE, Heimst€adt S, et al. Two PABPC1-binding sites in GW182 proteins promote miRNA-mediated gene silencing. EMBO J. 2010;29:4146–60. doi:10.1038/emboj.2010.274

[69] Kononenko A V., Mitkevich VA, Atkinson GC, et al. GTP-dependent structural rearrangement of the eRF1:eRF3 complex and eRF3 sequence motifs essential for PABP binding. Nucleic Acids Res. 2010;38:548–58. doi:10.1093/nar/gkp908

[70] Gallie DR, Liu R. Phylogenetic analysis reveals dynamic evolution of the poly(A)-binding protein gene family in plants. BMC Evol Biol. 2014;14:238. doi:10.1186/s12862-014-0238-4

[71] Derry MC, Yanagiya A, Martineau Y, et al. Regulation of poly (A)-binding protein through PABP-interacting proteins. Cold Spring Harb Symp Quant Biol. 2006;71:537–43. doi:10.1101/sqb.2006.71.061

[72] Minia I, Clayton C. Regulating a post-transcriptional regulator: protein phosphorylation, degradation and translational blockage in control of the trypanosome stress-response RNA-Binding Pro-tein ZC3H11. PLoS Pathog. 2016;12:1–31. doi:10.1371/journal.ppat.1005514

[73] Papadopoulou B, Roy G, Ouellette M. A novel antifolate resistance gene on the amplified H circle of Leishmania. EMBO J. 1992;11:3601–8

[74] McNicoll F, Muller M, Cloutier S, et al. Distinct 3 ^{0}-untranslated region elements regulate stage-specific mRNA accumulation and translation in Leishmania. J Biol Chem. 2005;280:35238–46. doi:10.1074/jbc.M507511200

[75] Cloutier S, Laverdiere M, Chou MN, et al. Translational control through eIF2alpha phosphorylation during the Leishmania differen-tiation process. PLoS One. 2012;7:e35085. doi:10.1371/journal. pone.0035085

[76] Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol. 1993;234:779–815. doi:10.1006/jmbi.1993.1626

[77] Leaver-Fay A, Tyka M, Lewis SM, et al. ROSETTA3: An object-ori-ented software suite for the simulation and design of macrom olecules. Methods Enzimol. 2011;487:545–74. doi:10.1016/B978-0-12-381270-4.00019-6

[78] Laskowski RA, MacArthur MW, Moss DS, et al. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr. 1993;26:283–91. doi:10.1107/S0021889892009944

[79] Padmanabhan PK, Samant M, Cloutier S, et al. Apoptosis-like programmed cell death induces antisense ribosomal RNA (rRNA) fragmentation and rRNA degradation in Leishmania. Cell Death Differ. 2012;19:1972–82. doi:10.1038/cdd.2012.85

[80] Mureev S, Kovtun O, Nguyen UTT, et al. Species-independent trans-lational leaders facilitate cell-free expression. Nat Biotechnol. 2009;27:747–52. doi:10.1038/nbt.1556

[81] Rezende AM, Assis LA, Nunes EC, et al. The translation initiation complex eIF3 in trypanosomatids and other pathogenic excavates–identification of conserved and divergent features based on ort hologue analysis. BMC Genomics. 2014;15:1175. doi:10.1186/1471-2164-15-1175